Exploring NADH:quinone oxidoreductases

Complex I and alternative NADH dehydrogenase (NDH-2)

Bruno Costa Marreiros

Dissertation presented to obtain the Ph.D degree in Biochemistry

Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa

Supervisor: Manuela M. Pereira

Oeiras, June, 2016

INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
ANTÓNIO XAVIER /UNL

Knowledge Creation

From left to right: José B. Pereira-Leal, Ana P. Batista, Volker Zickermann, Bruno C. Marreiros, Manuela M. Pereira, Cecília M. Arraiano and Teresa Catarino.

Oeiras, 21st June 2016

Acknowledgements

I would like to thank all the people that supported and helped me during these last years:

Dr. Manuela M. Pereira, my supervisor, for her support, dedication, enthusiasm and positive attitude, guidance, scientific discussions and friendship. For believe and bet in me to pursue the PhD.

Prof. Miguel Teixeira, for his support, advices, interest towards my work and for sharing his knowledge, ethics and pursue for perfectionism.

Dr. Ana Paula Batista, for her friendship, guidance, support, enthusiasm and for patience in teaching me and sharing all her knowledge and expertise. For all the help, advises and endless scientific discussions.

Dr. Ricardo O. Louro, for his interest towards my work, scientific discussions and advices.

Dr. Smilja Todorovic for scientific discussions in the Unit's meetings.

João Carita, for be always helpful and for growing cells.

Prof. Miguel Teixeira, Prof. Claudio Soares, Dr. Ricardo O. Louro, Dr. Inês Pereira, Dr. Teresa Catarino, Dr. Patricia N. Refojo, Dr. Afonso Duarte and Filipa Sena for critical reading of manuscripts.

Although the work performed in Max-Planck Institute for Biophysics, Frankfurt am Main was not included in this thesis, I would like to acknowledge Prof. Klaus Fendler and co-workers for receiving me at his group and teaching me Solid Supported Membrane based Electrophysiology to investigate ion translocating by complex I.

All my colleagues, former members and friends at the Metalloproteins and Bioenergetics Unit: Célia Romão, Ana Paula Batista, Patricia Refojo, Vera Gonçalves, Filipa Sousa, Afonso Duarte, Filipa Sena, Miguel Ribeiro, Rodrigo David, Elísio Silva, Filipe Sousa, Gabriel Martins, Sandra Santos, Cecilia Miranda, Filipa Calisto, Paulo Castro, Joana Carrilho, Joana Simões, Dalila Fernandes, Andreia Silva, Ana Filipa Pinto, Lara Paulo, Liliana Pinto, Guilherme Benedito and Martina Compagnucci.

All my colleagues, former colleagues and friends from Dr. Ana Paula Batista, Dr. Cláudio Gomes, Dr. Ricardo Louro, Dr. Lígia Saraiva and Dr. Inês Pereira groups.

All my colleagues and friends from the PhD program classes, 4th ITQB PhD meeting committee and ITQB Futsal games for helpful and friendly environment.

Ana Paula Batista, Bárbara Henriques, Gabriel Martins, Manuela Pereira, Miguel Ribeiro, Elísio Silva, Filipa Sena, Filipe Sousa, Patricia Refojo, Rodrigo David and Vera Gonçalves for friendshhip, support and great moments.

All my friends outside the ITQB, especially to Migós, for friendship and great adventures.

My family, especially to my parents, brother, sister and grandparents.

This work was financially supported by: Project LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) and by national funds through FCT - Fundação para a Ciência e a Tecnologia". The NMR spectrometers are part of The National NMR Facility, supported by Fundação para a Ciência e a Tecnologia (RECI/BBB-BQB/0230/2012). Fundação para a Ciência e a Tecnologia is acknowledge for financial support for funding PTDC/QUI-BIQ/100302/2008, Pest-OE/EQB/LA0004/2011, PTDC/BBB-BQB/2294/2012 and UID/CBQ/04612/2013.

Thesis publications

This Thesis is based on the following publications:

- Batista AP, **Marreiros BC**, Pereira MM (2011) “Decoupling of catalytic and transports activities of complex I from *Rhodothermus marinus* by a sodium/proton antiporter inhibitor”. ACS Chemical Biology 6 (5), 477-483. DOI: 10.1021/cb100380y.

- Batista AP*, **Marreiros BC***, Pereira MM (2012), *Equally Contributing Authors. “The role of proton and sodium ions in energy transduction by respiratory Complex I”. IUBMB life 64 (6), 492-498. DOI: 10.1002/iub.1050.

- **Marreiros BC**, Batista AP, Duarte AMS, Pereira MM (2013) “A missing link between complex I and group 4 membrane-bound [NiFe]-hydrogenases”. Biochimica *et* Biophysica Acta -Bioenergetics 1827 (2): 198-209. DOI: 10.1016/j.bbabio.2012.09.012.

- Batista AP, **Marreiros BC**, Pereira MM (2013) “The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes”. Biological chemistry 394 (5), 659-666. DOI: 10.1515/hsz-2012-0342.

- **Marreiros BC**, Batista AP, Pereira MM (2014) "Respiratory complex I from Escherichia coli does not transport Na^+ in the absence of its NuoL subunit" FEBS Letters, 588, 23, 4520-4525. DOI:10.1016/j.febslet.2014.10.030.

- Castro PJ, Silva AF, **Marreiros BC**, Batista AP, Pereira MM (2016), “Respiratory complex I: A dual relation with H^+ and Na^+?”. Biochimica et Biophysica Acta - Bioenergetics. DOI: 10.1016/j.bbabio.2015.12.008.

- **Marreiros BC**, Calisto F, Castro PJ, Duarte AM, Sena FV, Silva AF, Sousa FM, Teixeira M, Refojo PN and Pereira MM (2016), "Exploring membrane respiratory chains". Biochimica et Biophysica Acta - Bioenergetics. DOI: 10.1016/j.bbabio.2016.03.028.

- **Marreiros BC**, Sena FV, Sousa FM, Batista AP, Pereira MM (2016) "Type II NADH:quinone oxidoreductase family: Phylogenetic distribution, Structural diversity and Evolutionary divergences". Environmental Microbiology. DOI: 10.1111/1462-2920.13352.

Publication not included in this thesis:

- Batista AP, **Marreiros BC**, Louro RO, Pereira MM (2012) "Study of ion translocation by respiratory complex I. A new insight using (23)Na NMR spectroscopy" Biochim Biophys Acta, 1817(10):1810-6. DOI: 10.1016/j.bbabio.2012.03.009
-
- .

Summary

The work presented in this Thesis addressed the diversity of respiratory membrane enzymes with special focus on type I and type II NADH:quinone oxidoreductases. The searching for common denominators that characterize these enzymes in order to find minimal functional elements relevant for their operation was the main rational in this work.

Type I NADH:quinone oxidoreductase (Complex I) is able to couple the redox reaction to ion transport across the membrane. The prokaryotic enzyme is generally composed of 14 core subunits, summing of approximately 550 kDa, arranged in an L-shaped structure. The arms of the L are constituted by a hydrophilic part with 7 subunits and a membrane part with the remaining 7 subunits. The hydrophilic part contains all cofactors (one FMN and 8 or 9 iron-sulfur centers) including the catalytic sites for NADH oxidation and quinone reduction. In brief, the electrons from NADH are conducted from FMN (in the top of the hydrophilic part) to the quinone binding site (near the membrane interface) through a series of iron-sulfur centers. The oxidoreduction reaction is exergonic and the corresponding energy is transduced probably through conformational changes at the membrane part to the establishment of a difference in electrochemical potential by ion transport across the membrane. The membrane part includes 3 subunits (NuoL, M and N) homologous to

each other and to *bona fide* Na^+/H^+ antiporters, of the so called Mrp family. NuoL is longer than the other two antiporters-like subunits due to the presence of a long amphipathic α-helix (HL) at the C-terminal which spans over the 3 antiporter-like subunits. Based on the crystallographic structure from Complex I, an ion pathway was proposed for each antiporter-like subunit.

"How does the coupling mechanism occur?" and "What is the nature of the ions transported?" were main questions of this thesis. To address these questions we searched for the common features of Complex I enzymes and related complexes in order to identify the common denominator and energy transduction unit of these enzymes. Using bioinformatics analysis for Complex I and all members of its family, including group 4 [NiFe] hydrogenases and energy-converting hydrogenase related complexes (Ehr, assigned to Complex I family in this thesis), we identified four subunits (soluble NuoB/D and transmembrane NuoH/antiporter-like subunits) that compose the core subunits necessary to couple the redox activity to the ion transport activity. Moreover, we identified that the antiporter-like subunit from the common denominator has a conserved structural element which is a long-amphipathic α-helix like the one present in NuoL subunits from Complex I. Our results defined for the first time the common denominator in Complex I family.

In order to obtain a deeper insight into the coupling mechanism and the nature of the ion used by Complex I, the enzymes from *Rhodothermus marinus* and *Escherichia coli* were studied. Using fluorescence and ^{23}Na-NMR spectroscopies H^+ and Na^+ transports

were monitored during NADH:menaquinone oxidoreductase activity of Complex I in the presence and absence of the inhibitor EIPA (5-N-ethyl-N-isopropyl-amiloride). We corroborated our previous results that showed Na^+ transport in the opposite way to that of the coupling ion, H^+, and we observed that the catalytic and transport activities could be decoupled at different concentrations of EIPA. These observations suggested the involvement of an indirect coupling mechanism, possibly through conformational changes, which was later supported by structural studies. We explored the role of NuoL subunit (with HL) in ion transport by Complex I using a functional Complex I devoid of NuoL subunit from *E. coli*. We observed that the mutated Complex I does not transport Na^+ but still transports H^+. These results were in agreement with the previous results obtained for Complex I from *R. marinus* showing that energy transduction is performed by two different processes, proton pumping and Na^+/H^+ antiporting; in addition we showed that the NuoL subunit has a role in the antiport activity.

Although the coupling mechanism of Complex I is still not fully clear, this work contributed to the understanding of the conserved elements involved in the coupling mechanism as well as of the nature of transported ion. Our results also opened new perspectives on the coupling mechanism by Complex I by studying other members of Complex I family that may have a similar operating mode.

Type II NADH:quinone oxidoreductase (NDH-2) has a molecular mass of approximately 45 kDa and contributes indirectly to the

establishment of the transmembrane potential by reducing quinones. This enzyme is characterized by having two dinucleotide binding domains to interact with FAD and NADH, and therefore belongs to the two-dinucleotide binding domain flavoprotein (tDBDF) superfamily. Besides these two domains, NDH-2 has a C-terminal domain composed of two amphipathic α-helices that interacts with the membrane in a monotopic manner. With the resolution of the crystallographic structures of NDH-2s from two bacteria and one fungi species the study of these proteins gained a new enthusiasm.

The two main questions addressed in this study were: “Are NDH-2s present in the three domains of life?” and “What are the conserved elements involved in the redox mechanism of these enzymes?”. To address these questions we identified the NDH-2 family among all the members of the tDBDF superfamily and further investigated the conserved elements within the NDH-2 family. Using bioinformatics analysis we observed that NDH-2s are present in the three domains of life, including *Homo sapiens*. This result may question the idea of using NDH-2 (present in several pathogenic species) as an attractive target for rational design of specific drugs.

We further addressed the amino acid conservation of NDH-2s and identified conserved structural elements that led us to propose a redox mechanism based on these conserved elements. The proposed mechanism may also give insights in the operating mode of other members of the tDBDF superfamily.

Finally, using the same rational described before for Complex I and NDH-2 we further explored the membrane-bound enzymes of the

respiratory chains in the three domains of life. Using bioinformatics analysis we obtained a thorough taxonomic profile of all the membrane respiratory enzymes which brought a novel and broad view on the diversity of the respiratory chains in the three domains of life. Our results contributed to a better understanding of the vast cellular metabolisms.

Sumário

O trabalho apresentado nesta tese foca a diversidade de enzimas membranares respiratórias, nomeadamente os enzimas NADH:quinona oxidorreductases do tipo I e II. Este trabalho teve como base a procura de denominadores comuns nestes enzimas a fim de encontrar os elementos funcionais indispensáveis ao seu modo de operação.

NADH:quinona oxidorredutase tipo I (Complexo I) acopla a reacção oxidação-redução ao transporte de iões através da membrana. O Complexo I procariótico é geralmente composto por 14 subunidades principais, dispostas em forma de 'L', e que perfazem uma massa molecular de aproximadamente 550 kDa.

Os braços do 'L' são constituídos por uma parte hidrofílica com 7 subunidades e uma parte membranar com as outras 7 subunidades. A parte hidrofílica contém todos os cofatores (um FMN e 8 ou 9 centros de ferro-enxofre), e é responsável pela oxidação do NADH e a redução da quinona. Sucintamente, dois electrões são transferidos para o FMN (situado no extremo do braço hidrofílico oposto à membrana) por oxidação do NADH. Estes electrões são posteriormente transferidos para o local de ligação da quinona (perto do braço membranar) através de uma série de centros de ferro-enxofre que formam um fio conductor ao longo do braço hidrofílico. A energia libertada por esta reacção exergónica origina provavelmente

alterações conformacionais no braço membranar induzindo o transporte de iões através da membrana e consequente formação e estabelecimento de um potencial electroquímico.

O braço membranar inclui subunidades (NuoL, M e N) homólogas entre si e aos *antiporters* de Na^+/H^+ pertencentes à familia *Mrp*. A subunidade NuoL é a mais longa devido à presença de uma longa helice α anfipática (HL) no seu C-terminal, a qual se estende ao longo das três subunidades semelhantes a *antiporters*. Com base na estrutura cristalográfica do Complexo I, um caminho para o transporte de iões foi proposto em cada uma das subunidades semelhantes a *antiporters*.

As principais questões abordadas neste estudo foram: "Como é que ocorre o mecanismo de acoplamento?" e "Qual a natureza/tipo dos iões transportados?" Com o intuito de responder a estas questões, procurámos os elementos comuns a todos os enzimas da familia do Complexo I com a finalidade de identificar o denominador comum para a transdução de energia. Assim, os Complexos I, os *group 4 [NiFe] membrane-bound hydrogenases* e os *energy-converting hydrogenase related complexes* (Ehr) (que são propostos nesta tese como um dos membros da família do Complexo I) foram analisados por metodologias bioinformáticas. Do estudo resultou a identificação de 4 subunidades que representam o denominador comum de toda a família do Complexo I. Estas 4 subunidades, 2 soluveis (NuoB e D) e 2 transmembranares (NuoH e uma semelhante a *antiporters*) compõem o modulo necessário para acoplar a reacção de oxidação-redução ao transporte de iões. Posteriormente, identificámos que a subunidade

semelhante a *antiporters* presente no denominador comum tem uma longa hélice α amfipática no C-terminal semelhante à presente na subunidade NuoL do Complexo I.

Os Complexos I de *Rhodothermus marinus* e *Escherichia coli* foram estudados a fim de contribuirmos para uma melhor compreensão do mecanismo de acoplamento e da natureza dos iões translocados pelo Complexo I. Usando espectroscopias de fluorescência e RMN de ^{23}Na, os transportes de H^+ e Na^+ foram monotorizados durante a actividade NADH:menaquinona oxidoreductase, do Complexo I na presenca e ausência do inibidor EIPA (5-(*N*-etil-*N*-isopropil)-amilorida). Os resultados obtidos mostraram que o Na^+ é transportado no sentido oposto ao ião de acoplamento (H^+) durante a actividade catalítica (corroborando resultados anteriores) e que ambas as actividades, catalítica e transporte, podem ser desacopladas usando diferentes concentrações de EIPA. Estas observações sugeriram a presença de um mecanismo de acoplamento indirecto possivelmente através de alterações conformacionais do braço membranar do Complexo I. Estas observações foram posteriormente suportadas com a resolução da estrutura cristalográfica do Complexo I.

Ao estudar um Complexo I, cataliticamente activo, de *E. coli* desprovido da subunidade NuoL (com HL), observamos que este enzima perde a actividade do transporte de Na^+ mas mantém parte do transporte de H^+. Estes resultados estão de acordo com os resultados obtidos para o Complexo I de *R. marinus*, pois mostram que a tradução de energia é realizada por dois processos diferentes:

transporte de H^+ e *antiporte* de Na^+/H^+; e que a subunidade NuoL tem um papel importante na actividade *antiporte* Na^+/H^+ no Complex I.

Embora o mecanismo de acoplamento do Complexo I ainda não seja totalmente claro, este trabalho contribui para uma melhor compreensão dos elementos envolvidos no acoplamento da actividade catalítica ao transporte de iões, bem como da natureza dos iões translocados. Estes resultados também abrem novas perspectivas para o estudo do Complexo I a partir dos outros membros da família, pois sugerem que todos têm um modo de funcionamento semelhante.

NADH:quinona oxidorredutase tipo II (NDH-2) tem uma massa molecular de aproximadamente 45 kDa e contribui indirectamente para o estabelecimento do potencial transmembranar pela redução de quinonas. Este enzima é caracterizado por ter dois domínios de ligação de dinucleótidos que interagem com o FAD e o NADH. NDH-2 pertence à superfamília dos *two-dinucleotide binding domain flavoprotein* (tDBDF). Para além destes dois domínios, a NDH-2 tem um domínio no C-terminal composto por duas hélices α anfipáticas que interagem com a membrana de uma forma monotópica. Com a resolução das estruturas cristalográficas de NDH-2s, de duas bactérias e de um fungo, o estudo destes enzimas ganhou um novo entusiasmo.

As principais questões abordadas no ambito desta tese foram: "Existem NDH-2s nos três domínios da vida?" e "Quais são os elementos conservados envolvidos no mecanismo de oxidação-redução destes enzimas?". Para responder a estas questões, identificamos as proteinas pertencentes à família das NDH-2s entre

todos os membros da superfamília tDBDF e investigamos os elementos conservados da família das NDH-2. Usando análises bioinformáticas observamos que as NDH-2s estão presentes nos três domínios da vida, incluindo *Homo sapiens*. Este resultado questiona a ideia de desenvolver medicamentos que têm como alvo a NDH-2 (nomeadamente a sua inibição) com vista a controlar diversas espécies patogénicas.

Com base na identifição de resíduos de aminoácidos e elementos estruturais conservados das NDH-2s, propusemos um mecanismo para estes enzimas. Sendo que este poderá também contribuir para uma melhor percepção do modo de funcionamento dos outros membros da superfamília tDBDF.

Finalmente, usando o mesmo raciocínio descrito anteriormente para o Complexo I e NDH-2, investigamos os enzimas membranares pertencentes a cadeias respiratórias para os três domínios de vida. Usando análises bioinformáticas obtivemos o perfil taxonómico de todos estes enzimas respiratórios que ilustra e fornece uma nova e ampla visão sobre a diversidade das cadeias respiratórias nos três domínios da vida. Estes resultados contribuem para uma melhor compreensão dos vastos metabolismos celulares.

Abbreviations

A. aeolicus: *Aquifex aeolicus*; ***A. ambivalens***: *Acidianus ambivalens*; ***A. ferrooxidans***: *Acidithiobacillus ferrooxidans*; ***A. hospitalis***: *Acidianus hospitalis*; ***A. tumefaciens***: *Agrobacterium tumefaciens*; ***A. vinosum***: *Allochromatium vinosum*; ***A. woodii***: *Acetobacterium woodii*; ***B. pseudofirmus***: *Bacillus pseudofirmus*; ***B. subtillis****: Bacillus subtillis*; ***B. burgdorferi***: *Borrelia burgdorferi*; ***B. taurus***: *Bovine taurus*; ***C. aurantiacus***: *Chloroflexus aurantiacus*; ***C. glutamicum***: *Corynebacterium glutamicum*; ***C. jejuni*:** *Campylobacter jejuni*; ***C. reinhardtii***: *Chlamydomonas reinhardtii*; ***C. thermarum***: *Caldalkalibacillus thermarum*; ***E. coli*:** *Escherichia coli*; ***F. oxysporum***: *Fusarium oxysporum*; ***G. oxydans*** : *Gluconobacter oxydans*; ***H. sapiens***: *Homo sapiens*; ***K. pneumonia***: *Klebsiella pneumonia*; ***N. crassa***: *Neurospora crassa*; ***M. mazei****: Methanosarcina mazei*; ***M. smegmatis***: *Mycobacterium smegmatis*; ***M. tuberculosis***: *Mycobacterium tuberculosis*; ***N. thermophilus****: Natranaerobius thermophilus*; ***P. denitrificans***: *Paracoccus denitrificans*; ***P. falciparum***: *Plasmodium falciparum*; ***P. furiosus****: Pyrococcus furiosus*; ***P. gingivalis***: *Porphyromonas gingivalis*; ***R. marinus***: *Rhodothermus marinus*; ***R. rubrum***: *Rhodospirillum rubrum*; ***S. aureus***: *Staphylococcus aureus*; ***S. cerevisiae***: *Saccharomyces cerevisiae*; ***T. brucei***: *Trypanosoma brucei* ; ***T. gondii***: *Toxoplasma gondii*; ***T. kodakarensis****: Thermococcus kodakarensis;* ***T. maritima*:** *Thermotoga maritima*; ***T. pallidum***: *Treponema pallidum*; ***T. pseudonana***: *Thalassiosira pseudonana*; ***T. thermophilus***: *Thermus thermophilus*; ***Y. lipolytica****: Yarrowia lipolytica*.

ΔΨ: membrane potential; **ΔμH$^+$**: proton electrochemical potential; **ΔE:** redox potential difference; **ΔpH**: H^+ concentration difference; **ΔpNa**: Na^+ concentration difference; **A**: absorvance; **ACIII**: alternative complex III; **ACMA**: 9-amino-6-chloro-2-methoxyacridine; **ADH**: alcohol:quinone oxidoreductase; **AIF**: apoptosis inducing factor; **AIF-M2/AMID**: apoptosis-inducing factor-homologous mitochondrion-associated inducer of death; **ALDH**: aldehyde:quinone oxidoreductase; **AMO**: ammonia monooxygenase; **AOX**: alternative quinol:oxygen oxidoreductase; **AprAB**:

adenosine phosphosulfate reductase; **ATP**: adenosine triphosphate; ***b*$_6$*f***: cytochrome b_6f complex; ***bc*$_1$**: cytochrome bc_1 complex; **BLAST**: basic local alignment search tool; **BN-PAGE:** blue native polyacrylamide gel electrophoresis; **C**: carbon atom; **CCCP**: carbonyl cyanide *m*-chlorophenyl hydrazine; **CH_3-H_4MPT**: methyltetrahydromethanopterin; **CoB-SH**: coenzyme B; **CoM-S-CH_3**: methyl-coenzyme M; **CoM-S–S-CoB**: heterodisulfide of coenzyme M and coenzyme B; **CoM-SH**: coenzyme M; **Complex I/CpI**: type I NADH:quinone oxidoreductase; **Cyt *c***: Cytochrome *c*; **cytochrome *bd* oxidase**: quinol:oxygen oxidoreductase; **DAADH**: D-amino acid:quinone oxidoreductase; **DDM**: n-dodecyl-β-D-maltoside; **deamino-NADH**: reduced nicotinamide hypoxanthine dinucleotide; **DHAP**: dihydroxyacetone phosphate; **DHODH**: dihidroorotate:quinone oxidoreductase; **DMN**: 2;3-dimethyl-1;4-naphthoquinone; **DMS**: dimethyl sulfide; **DmsABC**: DMSO reductase; **DMSO**: dimethyl sulfoxide; **DUF**: domain of unknown function; **Dsr**: dissimilatory sulfite-reductase complex; **DUQ**: decylubiquinone; **E-value**: expect value; **EIPA**: 5-(*N*-ethyl-*N*-isopropyl)-amiloride; **Ech**: energy-converting hydrogenase; **Eha**: energy-converting hydrogenase A; **Ehb**: energy-converting hydrogenase B; **Ehr**: energy-converting hydrogenase related complex; **EPR:** electron paramagnetic resonance; **ETF-QO**: electron transfer flavoprotein:quinone oxidoreductase; **FAD**: flavin adenine dinucleotide; **Fd**: ferredoxin; **Fdn-N**: formate:quinone oxidoreductase; **FHL-1**: formate hydrogen lyase 1; **Fpo**: $F_{420}H_2$:phenazine oxidoreductase; **G3P**: glycerol-3-phosphate; **GADH**: D-gluconate:quinone oxidoreductase; **G3PDH-GlpABC**: anaerobic glycerol-3-phosphate:quinone oxidoreductase; **G3PDH-GlpD**: aerobic glycerol-3-phosphate:quinone oxidoreductase; **GLDH**: glycerol:quinone oxidoreductase; **Group-1 [NiFe] Hydrogenase**: hydrogen:quinone oxidoreductase; **H_4MPT**: tetrahydromethanopterin; **HAR**: hexammineruthenium (III) chloride; **HCO**: Heme-copper Oxygen reductase; **HDR**: methanophenazine:heterodisulfide oxidoreductase; **HiPIP**: high potential iron-sulfur protein; **HL**: long amphipathic α-helix; **KCN**: potassium cyanide; **KEGG**: Kyoto encyclopedia of genes and genomes; **LQO**: lactate:quinone oxidoreductase; **methylenephosphonate**: thulium (III) 1;4;7;10-tetraazacyclododecane-1;4;7;10-tetrakis; **Mbh**: membrane bound hydrogenase; **Mbx**: membrane bound hydrogenase related complex; **mGDH**: glucose:quinone

oxidoreductase; **ML**: maximum-likelihood; **mPPase**: membrane-bound pyrophosphatase; **MQ**: menaquinone; **MQO**: malate:quinone oxidoreductase; **Mrp**: multiple resistence to pH; **MSA**: multiple sequence alignment; **Mtr**: methyltetrahydromethanopterin:coenzyme M methyltransferase; **N**: nitrogen atom; **N-side**: prokaryotic cytoplasm or mitochondrial matrix; **NAD(P)H**: nicotinamide adenine dinucleotide (Phosphate); **Nap**: periplasmic quinol:nitrate oxidoreductase; **Nar**: quinol:nitrate oxidoreductase; **NBT**: nitroblue tetrazolium; **ndh**: NADH dehydrogenase; **NDH-2**: type II NADH:quinone oxidoreductase; **Nhc**: cytochrome:quinone oxidoreductase; **NJ**: neighbor-joining; **NOR**: nitric oxide reductase; **Nqo**: NADH:quinone oxidoreductase; **NQR**: Na^+-translocating NADH:quinone oxidoreductase; **NrfABCD**: quinol:nitrite oxidoreductase; **NrfHA**: quinol:nitrite oxidoreductase; **Nuo**: NADH:quinone oxidoreductase; **O**: oxygen atom; **OAD**: Na^+-translocating decarboxylase; **Oxonol VI**: 1;5-Bis (5-oxo-3-propylisoxazol-4-yl) pentamethine oxonol; **P**: phosphorus atom; **P-side**: prokaryotic periplasm or mitochondrial intermembrane space; **P5C**: (S)-1-pyrroline-5-carboxylate; **PDB**: protein data bank; **Pfam**: database of protein families; **Pi:** inorganic phosphate; **pMMO**: particulate methane monooxygenase; **Pnt**: membrane-bound transhydrogenase; **PPi:** inorganic pyrophosphate; **PPOR**: protoporphyrinogen IX:quinone oxidoreductase; **PQ**: plastoquinone; **PQO**: pyruvate:quinone oxidoreductase; **PRODH**: L-proline:quinone oxidoreductase; **PSII**: photosystem II; **PsrABC**: polysulfide reductase; **QRC**: cytochrome c_3:quinone oxidoreductase; **Qmo**: quinol:electron acceptor oxidoreductase; **Rnf**: ferredoxin:methanophenazine/NAD^+ oxidoreductase; **SDH**: succinate:quinone oxidoreductase; **SLDH**: D-sorbitol:quinone oxidoreductase; **SQR**: sulfide:quinone oxidoreductase; **SreABC**: sulfur redutase complex; **tDBDF:** two-dinucleotide binding domains flavoproteins; **TEMPO**: 2;2;6;6-tetramethyl-1-piperidinylox; **TM**: transmembrane; **TMA**: trimethylamine; **TMAO reductase**: quinol:trimethylamine N-oxide oxidoreductase; **Tpl c_3**: type I cytochrome c_3; **TQO reductase**: thiosulfite:quinone oxidoreductase; **TtrABC**: tetrathionate reductase, **UniProt**: universal protein resource; **UQ**: ubiquinone; **VBA**: visual basic for applications; **w/v**: weight/volume.

Table of Contents

Part I - Introduction

Chapter I - Exploring energy transduction by respiratory enzymes

Part II - Results

Chapter II – Complex I family

Chapter III – NDH-2 family

Chapter IV - Exploring membrane respiratory chains

Part III - Final discussion

Chapter V – Final discussion

Part I

Introduction

Chapter I

Exploring energy transduction by respiratory enzymes

Chapter I - Exploring energy transduction by respiratory enzymes

I.1 - Energy transduction by respiratory enzymes

This section is based on the following publications:

Batista AP*, **Marreiros BC***, Pereira MM (2012), "The role of proton and sodium ions in energy transduction by respiratory Complex I" IUBMB life 64 (6), 492-498. DOI: 10.1002/iub.1050. *Equally Contributing Authors.

Castro PJ, Silva AF, **Marreiros BC**, Batista AP, Pereira MM (2016), "Respiratory complex I: A dual relation with H^+ and Na^+?" Biochimica et Biophysica Acta - Bioenergetics. DOI:10.1016/j.bbabio.2015.12.008.

Marreiros BC, Calisto F, Castro PJ, Duarte AM, Sena FV, Silva AF, Sousa FM, Teixeira M, Refojo PN and Pereira MM (2016), "Exploring membrane respiratory chains" Biochimica et Biophysica Acta Bioenergetics. DOI: 10.1016/j.bbabio.2016.03.028.

I.1 - Energy transduction by respiratory enzymes

I.1.1 Energy transduction is a key process for life

In cells, most of the energy is transduced by membrane proteins, present in electron transfer chains, during cellular respiration or photosynthesis. The exergonic electron transfer from electron donors to electron acceptors is coupled to charge translocation across the membrane, establishing a transmembrane difference of electrochemical potential, as proposed in the Chemiosmotic Theory [1]. This potential is then used for endergonic processes such as ATP synthesis, active transport or motility.

I.1.2 Transmembrane difference of electrochemical potential is vital for life

The Chemiosmotic theory, formulated in 1961, considered the establishment of a proton electrochemical potential across the inner mitochondrial and plastidial membrane(s) ($\Delta\mu H^+$ in $kJ.mol^{-1}$ units, corresponding to proton motive force, *pmf* or Δp, in mV units), generated by the proton translocation activity of membrane bound respiratory complexes, which has an electrical (Ψ) and a chemical component (namely, pH) [1].

It was postulated that light energy, as well as respiration energy are transduced into a $\Delta\mu H^+$, which is then used for ATP synthesis, solute transport and motility. Later, this concept was

experimentally verified by showing that i) respiratory and photosynthetic complexes were directly involved in $\Delta\mu H^+$ formation, ii) ATP synthesis could be supported by an artificially-imposed $\Delta\mu H^+$ and iii) discharge of $\Delta\mu H^+$ abolished ATP synthesis, solute transport and motility [2].

In 1968, Mitchell proposed that ΔpH, built by the respiratory chain, could be used by Na^+/H^+ antiporters in order to create a difference in Na^+ concentration across the cristae membrane of Rat Liver Mitochondria (ΔpNa) [3-5]. Moreover, the transport of Na^+ across the membrane was shown to help to maintain $\Delta\mu H^+$ and pH homeostasis in the case of limitation of substrates or growth in alkaline environments [6, 7]. Since then, Na^+/H^+ antiporters have been widely described in prokaryotes and eukaryotes, establishing the foundations for the study of the bioenergetic role of Na^+ electrochemical potential [8-10]. Widespread in Bacteria and Archaea domains are organisms that possess primary Na^+ pumps [11]. The first evidence for the role of Na^+ as a coupling ion emerged when a primary Na^+ pump activity was determined for the oxaloacetate decarboxylase of *Klebsiella aerogenes* [12]. Additionally, it was observed that in alkaline conditions, NADH:quinone oxidoreductase activity of the respiratory chain of the marine bacterium *Vibrio alginolyticus* is coupled to Na^+ translocation across the membrane [13, 14]. Furthermore, it was recognized that some extremophilic (hyperthermophilic, alkaliphilic and/or halophilic) organisms and bacterial pathogens seem to depend on Na^+ to survive and grow [15-17]. In fact, it was demonstrated that, in some circumstances, Na^+ can

substitute H^+ as coupling ion and the resulting $\Delta\mu Na^+$ can be used for ATP synthesis, motility and/or solute import [12, 14, 18, 19]. This is possible because, from a thermodynamic point of view, the $\Delta\mu H^+$ and $\Delta\mu Na^+$ are equivalent, and composed of both chemical, ΔpH or ΔpNa, and electrical components, ΔΨ [11, 15].

The energy available in an ion electrochemical potential can be obtained taking into account both chemical and electrical components differences between the two sides of a membrane (ΔΨ). The ΔG involved in the transport of a cation (*x,* of valence *m*) from a negative (A) to a positive (B) side of the membrane, and against both a concentration gradient and difference of electrical potential, is described by,

$$\Delta\mu_{x^{m+}} = \Delta G = RT\ln\left(\frac{[x^+]_B}{[x^+]_A}\right) + mF\Delta\Psi \qquad \text{(Equation 1)}$$

where *F* is the Faraday constant (96.5 $kJ.V^{-1}.mol^{-1}$).

Most organisms use both H^+ and Na^+ cycles, either alternatively or concomitantly, and several enzymes seem to participate in the two cycles, being able to use either Na^+ and/or H^+ ions [2, 20-28]. The interconversion between $\Delta\mu Na^+$ and $\Delta\mu H^+$ may increase the robustness and adaptation of organisms to different environments and stress conditions, as observed in bacterial pathogens, for example. These organisms need to survive in their natural environment and once inside the host a rapid adaptation to the new conditions is required. In fact, the coexistence of H^+ and Na^+ cycles in most bacterial

microorganisms is guaranteed by Na^+ and H^+ transporters in respiratory chains and multiple Na^+/H^+ antiporters [17].

While in prokaryotes both Na^+ and H^+ cycles seem to coexist in the same membrane, in animal cells, H^+ and Na^+ cycles occur simultaneously, but in different cell locations. On the one hand, the establishment of an exclusive proton electrochemical potential across the inner mitochondrial membrane, by the electron transport chain, is essential for respiration. On the other hand, in the cytoplasmic membranes of most animal cells, a Na^+ cycle is established. The electrogenic Na^+/K^+-ATPase, which is present in the cytoplasmic membrane of all animals, uses the energy released by the hydrolysis of ATP to exchange 3 Na^+ ions with 2 K^+ ions, both against their concentration gradients [29, 30]. The established Na^+ electrochemical potential provides the driving force for many antiporters, such as Na^+/H^+ antiporter and Na^+/Ca^{2+} antiporter [31, 32]. The ionic fluxes promoted by the Na^+ cycle are involved in several cellular processes, such as the regulation of cell volume, the maintenance of the cell membrane resting potential, the generation of action potential in neuronal cells, the regulation of signal transduction mechanisms and the import of glucose, amino acids and other nutrients into the cell by processes that require a ΔpNa [33-35].

In conclusion, the bioenergetics of the cell is determined by the global operation of cycles of the different ions, namely H^+ and Na^+, which allow the establishment and maintenance of the transmembrane difference of electrochemical potential. Several enzymes are involved in both cycles or exclusively in only one.

I.1.3 Membrane respiratory enzymes

Acquisition of energy is central to life. Organisms need energy for the establishment and maintenance of a transmembrane difference in electrochemical potential allowing cells to import and export metabolites or to their motility. The membrane potential is established by a variety of membrane bound respiratory complexes that mostly are able to couple chemical reactions to ion translocation across the membrane.

In vertebrates, energy is generated in mitochondria by the process of oxidative phosphorylation. During the oxidation of nutrients such as glucose, amino acids or fatty acids, reduced metabolites with low reduction potentials, such as NADH or succinate, are produced in the mitochondrial matrix. These reduced metabolites provide electrons to the respiratory chain, placed in the inner mitochondrial membrane, which are conducted through several electron carriers and protein complexes containing redox centers with progressively higher reduction potential, until the final electron acceptor, O_2. The energy stored by this electrochemical potential is subsequently used by ATP synthase for the synthesis of adenosine triphosphate (ATP), by transhydrogenase for the reduction of $NADP^+$ by NADH and by transporters to import or export metabolites.

The membrane bound respiratory complexes were proposed to have different structural and functional organizations based on two models. The random state model suggests that the membrane bound

respiratory enzymes are randomly distributed in the membrane, with the electron transfer occurring based on the diffusion rates and collisional interactions of the electron carriers. On the contrary, the solid state model suggests that the membrane bound respiratory enzymes are structural and functional organized in supercomplexes, which may improve the oxidative phosphorylation processes by stabilization of individual complexes, channeling of the electron carriers enhancing the electron transfer rates and regulation of the formation of reactive oxygen species [36, 37].

I.1.3.1 Mitochondrial respiratory chains in mammals

The canonical respiratory chain of mammals consists of four multisubunit protein complexes, Complexes I, II, III and IV and of the mobile electron carriers ubiquinone and cytochrome *c* (Figure I.1.1). Electrons resultant from the oxidation of NADH enter the respiratory chain at the level of Complex I (NADH:ubiquinone oxidoreductase).

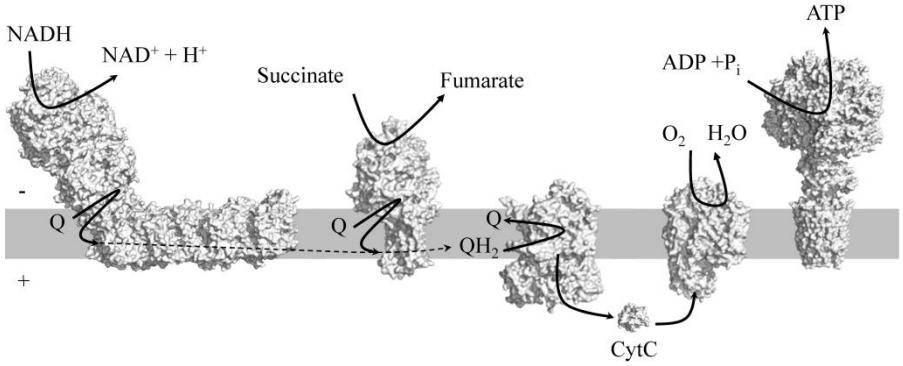


Figure I.1.1: Mitochondrial respiratory chain. Schematic representation of the classic mitochondrial respiratory complexes. The electron donors, NADH and Succinate are oxidized by complexes I and II, respectively, with concomitant

reduction of quinone to quinol. Complex III oxidizes quinol and reduces cytochrome *c*. Complex IV oxidizes cytochrome *c* and reduces oxygen to water. Complexes I, III and IV contribute to the establishment of the transmembrane difference in electrochemical potential by translocating protons across the membrane. Complex V uses the membrane potential for the synthesis of ATP (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively). The complexes's structures were adapted from: Complex I (PDB:4HEA); complex II (PDB:1ZOY); Complex III (PDB:1ZRT); Cytochrome *c* (PDB:4N0K); Complex IV (PDB:2YEV); Complex V (PDB:2XOK).

This enzyme catalyzes the two-electron oxidation of NADH and the reduction of ubiquinone to ubiquinol, coupled to the translocation of possibly four protons from the mitochondrial matrix to the intermembrane space, contributing in this way to the establishment of the membrane potential. Mitochondrial Complex I consists of 14 core subunits and 30 accessory subunits [38]. Complex II (succinate:quinone oxidoreductase or succinate dehydrogenase, SDH) is constituted by four subunits and catalyzes the two-electrons oxidation of succinate to fumarate with the reduction of ubiquinone to ubiquinol [39]. This reaction is not coupled to charge translocation. Besides Complexes I and II, at least three other single subunit enzymes in mammalian mitochondria feed electrons to the quinone/quinol pool without involving proton translocation across the membrane: Electron transfer flavoprotein:ubiquinone oxidoreductase (ETF-QO), which mediates the electron transfer between different mitochondrial flavoprotein dehydrogenases and the quinone/quinol pool [40]; glycerol-3-phosphate dehydrogenase (G3PDH), which catalyzes the oxidation of glycerol-3-phosphate, and dihydroorotate dehydrogenase (DHODH), involved in pyrimidine biosynthesis [41]. Complex III (ubiquinol:cytochrome *c* oxidoreductase or cytochrome bc_1 complex)

catalyzes the electron transfer from ubiquinol to cytochrome *c*, coupled to the translocation of two protons per oxidized ubiquinol across the inner mitochondrial membrane, contributing directly to the generation of the membrane potential [41]. In mitochondria of mammals, Complex III contains 11 subunits [42]. The final step in the mammalian mitochondrial respiratory chain is the transfer of electrons from reduced cytochrome *c* to O_2, forming 2 H_2O molecules. This four-electron reaction is catalyzed by Complex IV (cytochrome *c* oxidase, or cytochrome *c*:oxygen oxidoreductase), which pumps $4H^+$ *per* 4 e^- across the membrane [43]. Mitochondrial complex IV is composed of two catalytic subunits and 11 additional subunits [41].

I.1.3.2 Non mammal mitochondrial respiratory chains: plants, fungi and protists

The classic electron transport chain in plant mitochondria is similar to that of animal mitochondria: it comprises the four core complexes, Complexes I, II, III and IV. Complex I is composed of up to 50 subunits, 17 of which are different from the mammalian homologue [44, 45]. One of those distinct components was identified as a carbonic anhydrase [46]. Complex II is composed of four subunits and other additional proteins of unknown function in plants [44]. Ubiquinol produced by Complexes I and II is oxidized by Complex III. [47]. Plant Complex IV also shows supplementary proteins to the catalytic subunits [44].

Plants electron transfer chain also includes an alternative oxidase (AOX), which catalyzes the reduction of oxygen to H_2O directly

from quinol oxidation and without proton translocation across the membrane. The expression of AOX is manifested by the ability of plants to respire in the presence of cyanide, a potent inhibitor of Complex IV [45]. Additionally, several dehydrogenases can directly transfer electrons to ubiquinone such as alternative NAD(P)H dehydrogenases (type II NADH:quinone oxidoreductase, NDH-2), G3PDH and ETF-QO [44, 45].

Mitochondrial electron transport chains of fungi and protozoa also contain protein complexes similar to those of the classic mammalian mitochondrial chain [48-50]. Most fungal mitochondria have the classic respiratory chain, but some fungi, such as *Saccharomyces cerevisiae*, lack Complex I [51]. Instead, these fungi have genes encoding NDH-2s [52]. Also, usually fungi have additional components, such as AOX [53, 54].

Fungal parasites, like obligate intracellular Microsporidians, lack genes for energy metabolism and are strictly dependent on the host for energy. They uptake ATP from their host cells via ADP/ATP translocases located in their plasma membranes [55].

Several filamentous fungi are able to survive in anoxic conditions, being capable to produce ATP using nitrogen or sulfur as final electron acceptor [56, 57]. *Fusarium oxysporum,* from the Ascomycetes phylum has the ability to reduce nitrate to nitrous oxide, in a process coupled to ATP synthesis in the mitochondria, which is catalyzed by nitrate reductase, nitrite reductase and nitric oxide reductase [56, 58]. *F. oxysporum* is also able to reduce inorganic sulfur to hydrogen sulfide by a NADH-dependent sulfur reductase [57]. Some

anaerobic fungi, such as *Neocallimastix patriciarum*, lack mitochondria. Instead, they possess hydrogen and ATP-generating organelles called hydrogenosomes [59].

Protists may experience different environments and are able to live as free living organisms, parasites or symbionts of multicellular eukaryotes. For example, the free living aerobic amoeboflagellate *Naegleria gruberi* possesses the canonical mitochondrial respiratory enzymes and in addition it has NDH-2 and AOX [60]. *Euglena gracilis,* a single-celled flagellate protist, has a typical eukaryotic mitochondria but includes lactate:quinone oxidoreductase (LQO) and AOX [61, 62]. In *Pygsuia biforma*, an amoeba-like organism, ETF-QO, G3PDH and AOX are components of its respiratory chain [63].

A number of pathogenic protist species from low oxygen habitats, including *Trichomonas vaginalis*, *Giardia intestinalis* and *Entamoeba histolytica*, lack typical mitochondria [64-67]. For example, *T. vaginalis* instead of mitochondria contains hydrogenosomes, which produce hydrogen as an end product of fermentative energy metabolism [68]. *G. intestinalis* has mitochondrial-like genes and possesses small double-membrane-bound organelles named mitosomes that do not synthesize ATP [69].

Life cycles of protozoan parasites include development in the vector and host species and the mitochondrial metabolism changes depending on the life cycle stage, switching between active and fully repressed states [70, 71]. The Apicomplexa, a group of obligate intracellular parasites, include the protists *Plasmodium, Toxoplasma* and *Cryptosporidium*, the causative agents for malaria, toxoplasmosis

and cryptosporidiosis, respectively. *Cryptosporidium* species have lost the respective mtDNA and most mitochondrial enzymes, with the exception of AOX, malate:quinone oxidoreductase (MQO) and NDH-2 [72].

The anaerobic ciliates protists *Nyctotherus ovalis* and *Blastocystis* sp., for example have anaerobically functioning mitochondria that uses protons and fumarate as final electron acceptors and produce hydrogen and succinate, respectively [73-75]. In these microorganisms, the electron transport chain is composed of Complex I that generates a membrane potential, and a quinol:fumarate oxidoreductase, but lacks Complexes III and IV [73, 74].

I.1.3.3 Chloroplastidial respiratory chain

Chloroplasts are organized in a similar way to that of mitochondria: a two membrane system composed by an outer and an inner membranes and a large space surrounded by the inner membrane, the stroma. However, this organelle possesses a third distinct membrane, the thylakoid membrane, where electron transport occurs with proton translocation across the membrane [44, 76]. The chloroplast membrane possess unique components such as the photosystems [76]. Photosystem II (PSII), present in the membranes of oxygenic photosynthetic organisms, performs a series of reactions leading to the oxidation of water to protons and molecular oxygen [77, 78]. Electrons are transferred from PSII through the quinone/quinol pool to cytochrome b_6f complex

(plastoquinol:plastocyanin oxidoreductase) and then to photosystem I (PSI). Ultimately, in the electron transport chain of chloroplasts, ferrodoxin:$NADP^+$ oxidoreductase (FNR) reduces $NADP^+$ and oxidizes ferredoxin [78]. Several studies also suggested that this flavoprotein is involved in cyclic electron flow around PSI. In this case, electrons are transferred from PSI to cytochrome b_6f complex with formation of membrane potential and ATP synthesis [78, 79].

Cyclic electron flow around PSI, which recycles electrons from ferrodoxin to plastoquinone, comprises two pathways: one mediated by proteins so-called proton gradient regulators, PGR5 and PGR5-like 1 and another mediated by a Complex I homologue or NDH-2 [80-82]. This Complex I homologue is more related to that of Cyanobacteria than to the mitochondrial or proteobacterial ones and seems to have ferredoxin:plastoquinone oxidoreductase activity. Chloroplastidial Complex I is constituted by the 11 subunits of the cyanobacterial complex (ndhA-ndhK), which miss the so-called input model (corresponding to nuoE, F and G in *Escherichia coli)* [83], and three additional subunits, NdhM-NdhO [80]. Another component of the chloroplastidial respiratory chain is a non-heme diiron containing plastid terminal oxidase (PTOX), which belongs to the same family of AOX and mediates electron flow from plastoquinol to O_2 [84].

I.1.3.4 Prokaryotic respiratory chains

Prokaryotes show remarkable diversity, flexibility and robustness on their energetic metabolism, which allow them to live or survive under most diverse conditions (see section IV.1). Prokaryotes

may acquire energy by photosynthesis, fermentation or respiration. They can use different organic or inorganic compounds as electron donors to obtain energy, and a multiplicity of electron acceptors. Some organisms are aerobic and use oxygen as the last electron acceptor, while others use iron, sulfur or nitrogen compounds, or organic metabolites, as electron acceptors (Figure I.1.2).

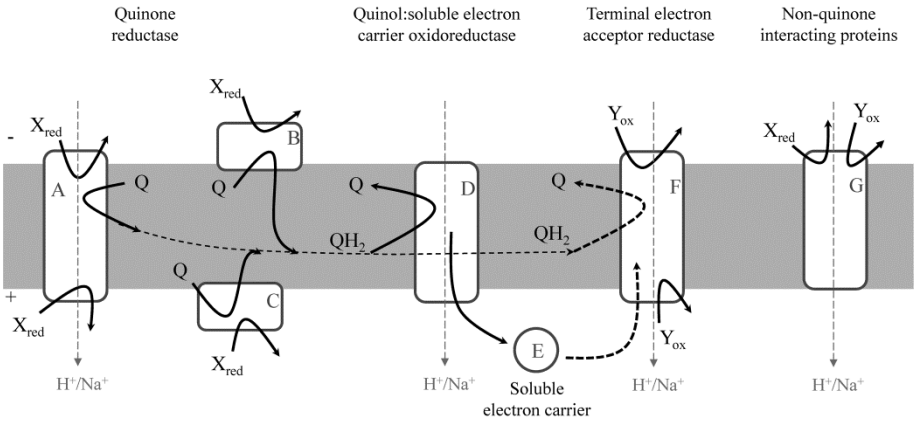


Figure I.1.2: Versatility of respiratory chains. Schematic representation of proteins and electron carriers involved in respiratory chains. A) transmembrane proteins with X:quinone oxidoreductase activity, with periplasmic or cytoplasmic exposed catalytic site. B) cytoplasmic facing monotopic proteins with X:quinone oxidoreductase activity. C) periplasmic facing monotopic proteins with X:quinone oxidoreductase activity. D) transmembrane proteins with quinol:soluble electron carrier oxidoreductase activity. E) soluble electron carrier that is reduced by protein D and is oxidized by protein F. F) transmembrane proteins with quinol:terminal electron acceptor or soluble electron carrier:terminal acceptor (Y) oxidoreductase activity. G) transmembrane proteins with oxidoreductase activity with no interaction with quinones. (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively).

I.1.4 Electron carriers: connection between respiratory complexes

Respiratory electron transfer is mediated by a group of electron carriers that can be in the membranes as the case of

quinones or can be soluble such as Cytochrome *c*, High potential iron-sulfur protein (HiPIP) and Type 1 copper proteins. They are capable of accepting and donating one or two electrons and perform functional connection between the respiratory membrane complexes. These electron carriers can function in a free state, diffusing between respiratory complexes, or in a trapped state, performing the electron transfer between respiratory complexes in a supercomplex assembly [37, 85].

I.1.4.1 Quinones

Isoprenoid quinones (commonly referred as quinones) are chemical entities that result from the fusion of a highly aliphatic isoprenylic chain with an aromatic ring system. They can be found in the membrane systems of cells, working as lipid soluble electron carriers playing a vital role in electron transport, oxidative stress control and gene regulation [86].

Isoprenoid quinones can be structurally divided according to their aromatic ring composition, into four main classes: benzoquinones, naphthoquinones, benzothiophenquinones and phenazines. Phenazines are not quinones from the organic chemistry point of view, but they are structurally related and perform a similar biochemical role in the cell membranes of some Archaea [87]. Each type of quinones can be further modified in the core aromatic ring composition, in addition to the variable length of the isoprenyl chain, generating an even greater structural diversity (Figure I.1.3) [88].

Benzoquinones can be divided in ubiquinones, plastoquinones and rhodoquinones; naphthoquinones can give rise to menaquinones, phyloquinones and thermoplasmaquinones; benzothiophenquinones can be modified into caldariellaquinones, sulfolobusquinones and tryciclicquinones; and phenazines are the structural precursors of methanophenazines (Figure I.1.3). Despite the structural variability, all quinones have in common the ability to be reduced giving rise to their quinol forms (or hydroquinones) in a reaction that involves two electrons and two protons.

Different substituents of the aromatic ring (CH_3 or NH_2 for example), or different ring structures, lead to different electron affinities of the quinone, translated into a change in the overall reduction potential of the molecule [89]. The examples of methanophenazine, -165 mV [90], caldariellaquinone, +100 mV [91], ubiquinone, +112 mV [92], rhodoquinone, -63 mV [93] and menaquinone, -74 mV [92], illustrate the interval of reduction potentials of quinones.

Some quinones have been described as taxonomic markers of a specific clade or metabolism (like caldariellaquinone for the extreme acidophilic Caldariella taxon [94, 95], or methanophenazine for methanogenic organisms [87]) meaning that they are taxa specific. Others like ubiquinone and menaquinone have been observed and characterized in a multitude of phyla.

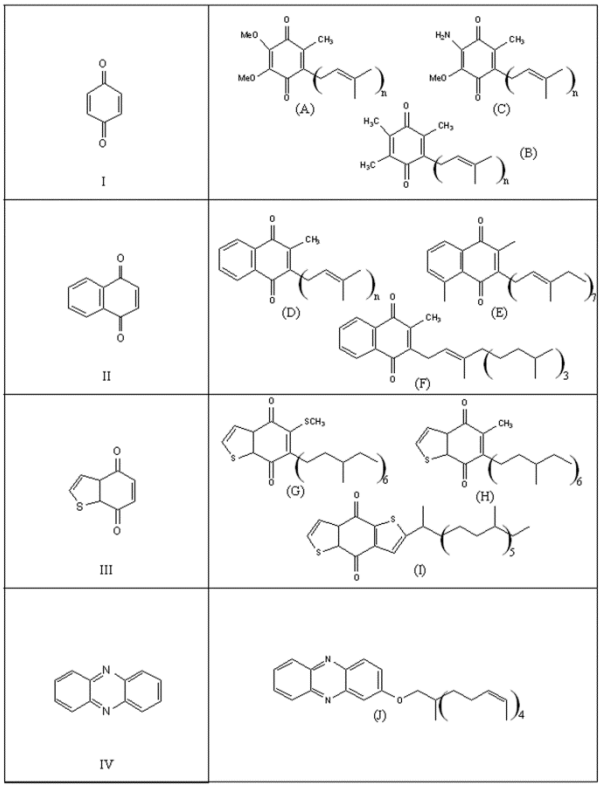


Figure I.1.3: Schematic representation of quinone structural precursors (I-IV) and physiological quinone structures (A-J). I) benzoquinone; II) naphtoquinone; III) benzothiophenquinone; IV) phenazine; A) ubiquinone; B)plastoquinone; C) rhodoquinone; D) menaquinone; E) thermoplasmaquinone; F) phylloquinone; G) caldariellaquinone; H) sulfolobusquinone; I) tryciclicquinone; J) methanophenazine.

I.1.4.2 'Soluble' electron carriers

Here, we describe the soluble electron carriers involved in the connections between respiratory proteins (Figure I.1.2). Cytochromes *c* are electron carriers having heme *c* as prosthetic group. The heme is covalently attached to the sulfydryls of a CXXCH motif via two α-thioether bonds [96, 97]. Cytochromes may connect different respiratory complexes, such as quinol:cytochrome *c* oxidoreductase and oxygen reductase, or be shuttles between membrane complexes

and soluble proteins or enzymes. Generally the monoheme cytochrome *c* involved in respiratory chains has a reduction potential of about +200 to + 300 mV [98].

HiPIP is a protein containing a single $[4Fe\text{-}4S]^{3+/2+}$ cluster with reduction potentials in the range of +50 to +450 mV, mainly found in purple photosynthetic bacteria [98-100]. HiPIPs have been shown to be electron donors to the reaction center in those bacteria [101, 102] and to the oxygen reductases of, for example, *Rhodothermus marinus* (a non-photosynthetic bacterium) [103] and *Rhodoferax fermentans* [104].

Type 1 copper proteins belong to a class named cupredoxins and are also frequently called blue copper proteins. The color of the proteins, blue (T1 copper) to green (perturbed T1 copper), depends on the presence and distance to the axial ligand (methionine or glutamate residues) of the copper center [105]. Over ten distinct sub-families were identified and characterized: Azurin, Amicyanin, Plastocyanin, Pseudoazurin, Rusticyanin, Auracyanin, Stellacyanin, Plantacyanin, Uclacyanin, Halocyanin, Sulfocyanin and Nitrosocyanin. These proteins, all electron carriers, were proposed to have roles in several physiological processes, including cellular response to redox stress, photosynthesis, Fe^{2+} oxidation, denitrification, besides O_2 or nitrite respiration [106, 107].

I.1.5 References

1. Mitchell, P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism, *Nature.* 191, 144-8.

2. Skulachev, V. P. (1991) Chemiosmotic systems in bioenergetics: H(+)-cycles and Na(+)-cycles, *Biosci Rep.* 11, 387-441; discussion 441-4.

3. Mitchell, P. & Moyle, J. (1969) Translocation of some anions cations and acids in rat liver mitochondria, *Eur J Biochem.* 9, 149-55.

4. Mitchell, P. & Moyle, J. (1967) Respiration-driven proton translocation in rat liver mitochondria, *Biochem J.* 105, 1147-62.

5. Mitchell, P. (1966) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation, *Biochim Biophys Acta.* 1807, 1507-38.

6. Brown, II, Galperin, M., Glagolev, A. N. & Skulachev, V. P. (1983) Utilization of energy stored in the form of Na+ and K+ ion gradients by bacterial cells, *Eur J Biochem.* 134, 345-9.

7. Ishikawa, T., Hama, H., Tsuda, M. & Tsuchiya, T. (1987) Isolation and properties of a mutant of Escherichia coli possessing defective Na+/H+ antiporter, *J Biol Chem.* 262, 7443-6.

8. West, I. C. & Mitchell, P. (1974) Proton/sodium ion antiport in *Escherichia coli*, *Biochem J.* 144, 87-90.

9. Krulwich, T. A. (1983) Na+/H+ antiporters, *Biochim Biophys Acta.* 726, 245-64.

10. Harold, F. M. & Papineau, D. (1972) Cation transport and electrogenesis by *Streptococcus faecalis*. II. Proton and sodium extrusion, *J Membr Biol.* 8, 45-62.

11. Skulachev, V. P. (1985) Membrane-linked energy transductions. Bioenergetic functions of sodium: H+ is not unique as a coupling ion, *Eur J Biochem.* 151, 199-208.

12. Dimroth, P. (1980) A new sodium-transport system energized by the decarboxylation of oxaloacetate, *FEBS Lett.* 122, 234-6.

13. Tokuda, H. & Unemoto, T. (1984) Na+ is translocated at NADH:quinone oxidoreductase segment in the respiratory chain of *Vibrio alginolyticus*, *J Biol Chem.* 259, 7785-90.

14. Tokuda, H. & Unemoto, T. (1981) A respiration-dependent primary sodium extrusion system functioning at alkaline pH in the marine bacterium *Vibrio alginolyticus*, *Biochem Biophys Res Commun.* 102, 265-71.

15. Lolkema, J. S., Speelmans, G. & Konings, W. N. (1994) Na(+)-coupled versus H(+)-coupled energy transduction in bacteria, *Biochim Biophys Acta.* 1187, 211-5.

16. Kogure, K. (1998) Bioenergetics of marine bacteria, *Curr Opin Biotechnol.* 9, 278-82.

17. Hase, C. C., Fedorova, N. D., Galperin, M. Y. & Dibrov, P. A. (2001) Sodium ion cycle in bacterial pathogens: evidence from cross-genome comparisons, *Microbiol Mol Biol Rev.* 65, 353-70, table of contents.

18. Skulachev, V. P. (1992) The laws of cell energetics, *Eur J Biochem.* 208, 203-9.

19. Skulachev, V. P. (1989) The sodium cycle: a novel type of bacterial energetics, *J Bioenerg Biomembr.* 21, 635-47.

20. Terahara, N., Krulwich, T. A. & Ito, M. (2008) Mutations alter the sodium versus proton use of a *Bacillus clausii* flagellar motor and confer dual ion use on *Bacillus subtilis* motors, *Proc Natl Acad Sci U S A.* 105, 14359-64.

21. Leone, V., Pogoryelov, D., Meier, T. & Faraldo-Gomez, J. D. (2015) On the principle of ion selectivity in Na+/H+-coupled membrane proteins: experimental and theoretical studies of an ATP synthase rotor, *Proc Natl Acad Sci U S A.* 112, E1057-66.

22. Luoto, H. H., Baykov, A. A., Lahti, R. & Malinen, A. M. (2013) Membrane-integral pyrophosphatase subfamily capable of translocating both Na+ and H+, *Proc Natl Acad Sci U S A.* 110, 1255-60.

23. Luoto, H. H., Nordbo, E., Baykov, A. A., Lahti, R. & Malinen, A. M. (2013) Membrane Na+-pyrophosphatases can transport protons at low sodium concentrations, *J Biol Chem.* 288, 35489-99.

24. Hama, H. & Wilson, T. H. (1992) Primary structure and characteristics of the melibiose carrier of *Klebsiella pneumoniae*, *J Biol Chem.* 267, 18371-6.

25. McMillan, D. G., Ferguson, S. A., Dey, D., Schroder, K., Aung, H. L., Carbone, V., Attwood, G. T., Ronimus, R. S., Meier, T., Janssen, P. H. & Cook, G. M. (2011) A1Ao-ATP synthase of *Methanobrevibacter ruminantium* couples sodium ions for ATP synthesis under physiological conditions, *J Biol Chem.* 286, 39882-92.

26. Laubinger, W. & Dimroth, P. (1989) The sodium ion translocating adenosinetriphosphatase of *Propionigenium modestum* pumps protons at low sodium ion concentrations, *Biochemistry.* 28, 7194-8.

27. Schlegel, K., Leone, V., Faraldo-Gomez, J. D. & Muller, V. (2012) Promiscuous archaeal ATP synthase concurrently coupled to Na+ and H+ translocation, *Proc Natl Acad Sci U S A.* 109, 947-52.

28. Jin, Y., Nair, A. & van Veen, H. W. (2014) Multidrug transport protein norM from *Vibrio cholerae* simultaneously couples to sodium- and proton-motive force, *J Biol Chem.* 289, 14624-32.

29. Skou, J. C. (1957) The influence of some cations on an adenosine triphosphatase from peripheral nerves, *Biochim Biophys Acta.* 23, 394-401.

30. Glynn, I. M. (1956) Sodium and potassium movements in human red cells, *J Physiol.* 134, 278-310.

31. Blaustein, M. P. & Lederer, W. J. (1999) Sodium/calcium exchange: its physiological implications, *Physiol Rev.* 79, 763-854.

32. Malo, M. E. & Fliegel, L. (2006) Physiological role and regulation of the Na+/H+ exchanger, *Can J Physiol Pharmacol.* 84, 1081-95.

33. Alvarez-Leefmans, F. J., Gamino, S. M. & Reuss, L. (1992) Cell volume changes upon sodium pump inhibition in Helix aspersa neurones, *J Physiol.* 458, 603-19.

34. Lodish H., B. A., Zipursky S.L., et al. (2000) *Cotransport by Symporters and Antiporters*, 4 edn, W. H. Freeman, New York.

35. Xie, Z. (2003) Molecular mechanisms of Na/K-ATPase-mediated signal transduction, *Ann N Y Acad Sci.* 986, 497-503.

36. Hackenbrock, C. R., Chazotte, B. & Gupte, S. S. (1986) The random collision model and a critical assessment of diffusion and collision in mitochondrial electron transport, *J Bioenerg Biomembr.* 18, 331-68.

37. Enriquez, J. A. (2016) Supramolecular Organization of Respiratory Complexes, *Annual review of physiology.* 78, 533-61.

38. Hirst, J. (2013) Mitochondrial Complex I, *Annu Rev Biochem.* 82, 551-75.

39. Lancaster, C. R. (2002) Succinate:quinone oxidoreductases: an overview, *Biochim Biophys Acta.* 1553, 1-6.

40. Watmough, N. J. & Frerman, F. E. (2010) The electron transfer flavoprotein: ubiquinone oxidoreductases, *Biochim Biophys Acta.* 1797, 1910-6.

41. Nicholls, D. G. & Ferguson, S. J. (2013) 5 - Respiratory chains in *Bioenergetics (Fourth Edition)* (Ferguson, D. G. N. J., ed) pp. 91-157, Academic Press, London.

42. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S. & Jap, B. K. (1998) Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex, *Science.* 281, 64-71.

43. Wikstrom, M. (1989) Identification of the electron transfers in cytochrome oxidase that are coupled to proton-pumping, *Nature.* 338, 776-8.

44. Millar, A. H., Whelan, J., Soole, K. L. & Day, D. A. (2011) Organization and regulation of mitochondrial respiration in plants, *Annual review of plant biology.* 62, 79-104.

45. Schertl, P. & Braun, H. P. (2014) Respiratory electron transfer pathways in plant mitochondria, *Frontiers in plant science.* 5, 163.

46. Sunderhaus, S., Dudkina, N. V., Jänsch, L., Klodmann, J., Heinemeyer, J., Perales, M., Zabaleta, E., Boekema, E. J. & Braun, H.-P. (2006) Carbonic Anhydrase Subunits Form a Matrix-exposed Domain Attached to the Membrane Arm of Mitochondrial Complex I in Plants, *J Biol Chem*.

47. (2015) *Plant Mitochondria | Frank Kempken | Springer*.

48. Abdrakhmanova, A., Zickermann, V., Bostina, M., Radermacher, M., Schagger, H., Kerscher, S. & Brandt, U. (2004) Subunit composition of mitochondrial Complex I from the yeast *Yarrowia lipolytica*, *Biochim Biophys Acta.* 1658, 148-56.

49. Marechal, A., Meunier, B., Lee, D., Orengo, C. & Rich, P. R. (2012) Yeast cytochrome *c* oxidase: a model system to study mitochondrial forms of the haem-copper oxidase superfamily, *Biochim Biophys Acta.* 1817, 620-8.

50. Gabaldon, T., Rainey, D. & Huynen, M. A. (2005) Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (Complex I), *Journal of molecular biology.* 348, 857-70.

51. Nosek, J. & Fukuhara, H. (1994) NADH dehydrogenase subunit genes in the mitochondrial DNA of yeasts, *Journal of bacteriology.* 176, 5622-30.

52. Luttik, M. A., Overkamp, K. M., Kotter, P., de Vries, S., van Dijken, J. P. & Pronk, J. T. (1998) The *Saccharomyces cerevisiae* NDE1 and NDE2 genes encode separate

mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH, *J Biol Chem.* 273, 24529-34.

53. Uribe, D. & Khachatourians, G. G. (2008) Identification and characterization of an alternative oxidase in the entomopathogenic fungus *Metarhizium anisopliae*, *Canadian journal of microbiology.* 54, 119-27.

54. Joseph-Horne, T., Hollomon, D. W. & Wood, P. M. (2001) Fungal respiration: a fusion of standard and alternative components, *Biochim Biophys Acta.* 1504, 179-95.

55. Tsaousis, A. D., Kunji, E. R., Goldberg, A. V., Lucocq, J. M., Hirt, R. P. & Embley, T. M. (2008) A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*, *Nature.* 453, 553-6.

56. Takaya, N., Kuwazaki, S., Adachi, Y., Suzuki, S., Kikuchi, T., Nakamura, H., Shiro, Y. & Shoun, H. (2003) Hybrid respiration in the denitrifying mitochondria of *Fusarium oxysporum*, *Journal of biochemistry.* 133, 461-5.

57. Abe, T., Hoshino, T., Nakamura, A. & Takaya, N. (2007) Anaerobic elemental sulfur reduction by fungus *Fusarium oxysporum*, *Bioscience, biotechnology, and biochemistry.* 71, 2402-7.

58. Kobayashi, M., Matsuo, Y., Takimoto, A., Suzuki, S., Maruo, F. & Shoun, H. (1996) Denitrification, a novel type of respiratory metabolism in fungal mitochondrion, *J Biol Chem.* 271, 16263-7.

59. van der Giezen, M., Sjollema, K. A., Artz, R. R., Alkema, W. & Prins, R. A. (1997) Hydrogenosomes in the anaerobic fungus *Neocallimastix frontalis* have a double membrane but lack an associated organelle genome, *FEBS Lett.* 408, 147-50.

60. Fritz-Laylin, L. K., Prochnik, S. E., Ginger, M. L., Dacks, J. B., Carpenter, M. L., Field, M. C., Kuo, A., Paredez, A., Chapman, J., Pham, J., Shu, S., Neupane, R., Cipriano, M., Mancuso, J., Tu, H., Salamov, A., Lindquist, E., Shapiro, H., Lucas, S., Grigoriev, I. V., Cande, W. Z., Fulton, C., Rokhsar, D. S. & Dawson, S. C. (2010) The genome of *Naegleria gruberi* illuminates early eukaryotic versatility, *Cell.* 140, 631-42.

61. Castro-Guerrero, N. A., Krab, K. & Moreno-Sanchez, R. (2004) The alternative respiratory pathway of euglena mitochondria, *J Bioenerg Biomembr.* 36, 459-69.

62. Moreno-Sanchez, R., Covian, R., Jasso-Chavez, R., Rodriguez-Enriquez, S., Pacheco-Moises, F. & Torres-Marquez, M. E. (2000) Oxidative phosphorylation supported by an alternative respiratory pathway in mitochondria from Euglena, *Biochim Biophys Acta.* 1457, 200-10.

63. Stairs, C. W., Eme, L., Brown, M. W., Mutsaers, C., Susko, E., Dellaire, G., Soanes, D. M., van der Giezen, M. & Roger, A. J. (2014) A SUF Fe-S cluster biogenesis system in the mitochondrion-related organelles of the anaerobic protist Pygsuia, *Current biology : CB.* 24, 1176-86.

64. Roger, A. J., Clark, C. G. & Doolittle, W. F. (1996) A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*, *Proc Natl Acad Sci U S A.* 93, 14618-22.

65. Roger, A. J., Svard, S. G., Tovar, J., Clark, C. G., Smith, M. W., Gillin, F. D. & Sogin, M. L. (1998) A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria, *Proc Natl Acad Sci U S A.* 95, 229-34.

66. Clark, C. G. & Roger, A. J. (1995) Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*, *Proc Natl Acad Sci U S A.* 92, 6518-21.

67. Stairs, C. W., Leger, M. M. & Roger, A. J. (2015) Diversity and origins of anaerobic metabolism in mitochondria and related organelles, *Philosophical transactions of the Royal Society of London Series B, Biological sciences.* 370, 20140326.

68. Muller, M. (1993) The hydrogenosome, *J Gen Microbiol.* 139, 2879-89.

69. Tovar, J., Leon-Avila, G., Sanchez, L. B., Sutak, R., Tachezy, J., van der Giezen, M., Hernandez, M., Muller, M. & Lucocq, J. M. (2003) Mitochondrial remnant organelles of Giardia function in iron-sulphur protein maturation, *Nature.* 426, 172-6.

70. Surve, S., Heestand, M., Panicucci, B., Schnaufer, A. & Parsons, M. (2012) Enigmatic presence of mitochondrial complex I in *Trypanosoma brucei* bloodstream forms, *Eukaryotic cell.* 11, 183-93.

71. Bringaud, F., Riviere, L. & Coustou, V. (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources, *Molecular and biochemical parasitology.* 149, 1-9.

72. Mogi, T. & Kita, K. (2010) Diversity in mitochondrial metabolic pathways in parasitic protists Plasmodium and Cryptosporidium, *Parasitology international.* 59, 305-12.

73. Boxma, B., de Graaf, R. M., van der Staay, G. W., van Alen, T. A., Ricard, G., Gabaldon, T., van Hoek, A. H., Moon-van der Staay, S. Y., Koopman, W. J., van Hellemond, J. J., Tielens, A. G., Friedrich, T., Veenhuis, M., Huynen, M. A. & Hackstein, J. H. (2005) An anaerobic mitochondrion that produces hydrogen, *Nature.* 434, 74-9.

74. de Graaf, R. M., Ricard, G., van Alen, T. A., Duarte, I., Dutilh, B. E., Burgtorf, C., Kuiper, J. W., van der Staay, G. W., Tielens, A. G., Huynen, M. A. & Hackstein, J. H. (2011) The organellar genome and metabolic potential of the hydrogen-producing mitochondrion of *Nyctotherus ovalis*, *Molecular biology and evolution.* 28, 2379-91.

75. Muller, M., Mentel, M., van Hellemond, J. J., Henze, K., Woehle, C., Gould, S. B., Yu, R. Y., van der Giezen, M., Tielens, A. G. & Martin, W. F. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes, *Microbiol Mol Biol Rev.* 76, 444-95.

76. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002) *Chloroplasts and Photosynthesis*, Garland Science.

77. Umena, Y., Kawakami, K., Shen, J.-R. & Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9[thinsp]A, *Nature.* 473, 55-60.

78. Rochaix, J. D. (2011) Reprint of: Regulation of photosynthetic electron transport ☆ ☆☆, *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 1807, 878–886.

79. Mulo, P. (2011) Chloroplast-targeted ferredoxin-$NADP^+$ oxidoreductase (FNR): Structure, function and location ☆, *Biochimica et Biophysica Acta (BBA) - Bioenergetics.* 1807, 927–934.

80. Peng, L., Yamamoto, H. & Shikanai, T. (2011) Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex, *Biochim Biophys Acta.* 1807, 945-53.

81. Jans, F., Mignolet, E., Houyoux, P.-A., Cardol, P., Ghysels, B., Cuiné, S., Cournac, L., Peltier, G., Remacle, C. & Franck, F. (2008) A type II NAD(P)H dehydrogenase mediates light-independent plastoquinone reduction in the chloroplast of Chlamydomonas, *Proc Natl Acad Sci U S A*.

82. Peltier, G., Aro, E. M. & Shikanai, T. (2015) NDH-1 and NDH-2 Plastoquinone Reductases in Oxygenic Photosynthesis, *Annual review of plant biology*.

83. Ifuku, K., Endo, T., Shikanai, T. & Aro, E. M. (2011) Structure of the chloroplast NADH dehydrogenase-like complex: nomenclature for nuclear-encoded subunits, *Plant Cell Physiol.* 52, 1560-8.

84. McDonald, A. E., Ivanov, A. G., Bode, R., Maxwell, D. P., Rodermel, S. R. & Huner, N. P. (2011) Flexibility in photosynthetic electron transport: the physiological role of plastoquinol terminal oxidase (PTOX), *Biochim Biophys Acta.* 1807, 954-67.

85. Trouillard, M., Meunier, B. & Rappaport, F. (2011) Questioning the functional relevance of mitochondrial supercomplexes by time-resolved analysis of the respiratory chain, *Proc Natl Acad Sci U S A.* 108, E1027-34.

86. Soballe, B. & Poole, R. K. (1999) Microbial ubiquinones: multiple roles in respiration, gene regulation and oxidative stress management, *Microbiology.* 145 (Pt 8), 1817-30.

87. Deppenmeier, U., Lienard, T. & Gottschalk, G. (1999) Novel reactions involved in energy conservation by methanogenic archaea, *FEBS Lett.* 457, 291-7.

88. Collins, M. D. & Jones, D. (1981) Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication, *Microbiol Rev.* 45, 316-54.

89. Conant, J. B. & Fieser, L. F. (1924) Reduction potentials of quinones II The potentials of certain derivatives of benzoquinone, naphthoquinone and anthraquinone, *J Am Chem Soc.* 46, 1858-1881.

90. Tietze, M., Beuchle, A., Lamla, I., Orth, N., Dehler, M., Greiner, G. & Beifuss, U. (2003) Redox potentials of methanophenazine and CoB-S-S-CoM, factors involved in electron transport in Methanogenic archaea, *Chembiochem : a European journal of chemical biology.* 4, 333-5.

91. Anemuller, S. & Schafer, G. (1990) Cytochrome-aa_3 from *Sulfolobus acidocaldarius* a Single-Subunit, Quinol-Oxidizing Archaebacterial Terminal Oxidase, *European Journal of Biochemistry.* 191, 297-305.

92. Thauer, R. K., Jungermann, K. & Decker, K. (1977) Energy conservation in chemotrophic anaerobic bacteria, *Bacteriol Rev.* 41, 100-80.

93. Van Hellemond, J. J., Klockiewicz, M., Gaasenbeek, C. P., Roos, M. H. & Tielens, A. G. (1995) Rhodoquinone and complex II of the electron transport chain in anaerobically functioning eukaryotes, *J Biol Chem.* 270, 31065-70.

94. De Rosa, M., Gambacorta, A., Millonig, G. & Bu'Lock, J. D. (1974) Convergent characters of extremely thermophilic acidophilic bacteria, *Experientia.* 30, 866-8.

95. De Rosa, M., De Rosa, S., Gambacorta, A. & Minale, L. (1977) Caldariellaquinone, a unique benzo(b)thiophen-4,7-quinone from *Caldariella acidophila*, an extremely thermophilic and acidophilic bacterium, *J Chem Soc Perkin 1*, 653-7.

96. Kranz, R., Lill, R., Goldman, B., Bonnard, G. & Merchant, S. (1998) Molecular mechanisms of cytochrome *c* biogenesis: three distinct systems, *Molecular microbiology.* 29, 383-96.

97. Salemme, F. R. (1977) Structure and function of cytochromes C, *Annu Rev Biochem.* 46, 299-329.

98. Meyer, T. E. & Cusanovich, M. A. (2003) Discovery and characterization of electron transfer proteins in the photosynthetic bacteria, *Photosynthesis research.* 76, 111-26.

99. Pereira, M. M., Carita, J. N. & Teixeira, M. (1999) Membrane-bound electron transfer chain of the thermohalophilic bacterium *Rhodothermus marinus*: characterization of the iron-sulfur centers from the dehydrogenases and investigation of the high-potential iron-sulfur protein function by in vitro reconstitution of the respiratory chain, *Biochemistry.* 38, 1276-83.

100. Meyer, T. E. (1994) Purification and properties of high-potential iron-sulfur proteins, *Methods in enzymology.* 243, 435-47.

101. Hochkoeppler, A., Zannoni, D., Ciurli, S., Meyer, T. E., Cusanovich, M. A. & Tollin, G. (1996) Kinetics of photo-induced electron transfer from high-potential iron-sulfur protein to the photosynthetic reaction center of the purple phototroph *Rhodoferax fermentans*, *Proc Natl Acad Sci U S A.* 93, 6998-7002.

102. Meyer, T. E., Bartsch, R. G., Cusanovich, M. A. & Tollin, G. (1993) Kinetics of photooxidation of soluble cytochromes, HiPIP, and azurin by the photosynthetic reaction center of the purple phototrophic bacterium *Rhodopseudomonas viridis*, *Biochemistry.* 32, 4719-26.

103. Santana, M., Pereira, M. M., Elias, N. P., Soares, C. M. & Teixeira, M. (2001) Gene cluster of *Rhodothermus marinus* high-potential iron-sulfur Protein: oxygen oxidoreductase, a caa_3-type oxidase belonging to the superfamily of heme-copper oxidases, *Journal of bacteriology.* 183, 687-99.

104. Hochkoeppler, A., Kofod, P. & Zannoni, D. (1995) HiPiP oxido-reductase activity in membranes from aerobically grown cells of the facultative phototroph *Rhodoferax fermentans*, *FEBS Lett.* 375, 197-200.

105. Wilson, T. D., Yu, Y. & Lu, Y. Understanding copper-thiolate containing electron transfer centers by incorporation of unnatural amino acids and the CuA center into the type 1 copper protein azurin. , *Coordination Chemistry Reviews.* 257, 260-276.

106. De Rienzo, F., Gabdoulline, R. R., Menziani, M. C. & Wade, R. C. (2000) Blue copper proteins: a comparative analysis of their molecular interaction properties, *Protein science : a publication of the Protein Society.* 9, 1439-54.

107. Dennison, C. (2005) Investigating the structure and function of cupredoxins, *Coordination Chemistry Reviews.* 249.

I.2 - Type I NADH:quinone oxidoreductase

This section is based on the following publications:

Batista AP*, **Marreiros BC***, Pereira MM (2012) "The role of proton and sodium ions in energy transduction by respiratory Complex I" IUBMB life 64 (6), 492-498. DOI: 10.1002/iub.1050. *Equally Contributing Authors.

Marreiros BC, Batista AP, Duarte AMS, Pereira MM (2013) "A missing link between complex I and group 4 membrane-bound [NiFe]-hydrogenases" Biochimica *et* Biophysica Acta (BBA)-Bioenergetics 1827 (2): 198-209. DOI: 10.1016/j.bbabio.2012.09.012

Castro PJ, Silva AF, **Marreiros BC**, Batista AP, Pereira MM (2016), "Respiratory complex I: A dual relation with H^+ and Na^+?" Biochimica *et* Biophysica Acta (BBA)-Bioenergetics. DOI: 10.1016/j.bbabio.2015.12.008

I.2 - Type I NADH:quinone oxidoreductase

Complex I (NADH:quinone oxidoreductase, EC. 1.6.5.3) plays a central role in the energy transduction processes of the respiratory chains from mitochondria and many bacteria and it dysfunction has been implicated in a wide range of pathologies such as Leigh's disease and neurodegenerative diseases [1]. Complex I functions as an entry point for electrons into those respiratory chains and catalyzes the two electron oxidation of NADH and the reduction of quinone, coupled to charge translocation from the negatively charged side (N-side, prokaryotic cytoplasm or mitochondrial matrix) to the positively charged side (P-side, prokaryotic periplasm or mitochondrial intermembrane space) of the membrane. This process results in the establishment of a transmembrane difference of electrochemical potential. It is an L-shaped membrane protein, consisting of a peripheral and a membrane arm. The peripheral arm contains a series of iron-sulfur centers and a FMN at the catalytic site, where NADH is oxidized [2]. The membrane arm includes the cation translocating machinery [3, 4].

Homologues of complex I exist in Bacteria, Archaea and Eukarya. In the latter domain, the complex is present both in mitochondria and chloroplasts. Bacterial complex I is generally composed of 14 subunits (~550 kDa) designated NuoA to N (from NADH:ubiquinone oxidoreductase) or as Nqo1 to 14 (from NADH:quinone oxidoreductase), whereas the mitochondrial enzyme is constituted by more than 40 subunits (~1 MDa) [5]. The electron donor

of several bacterial and mitochondrial complexes is NADH, while some archaeal complex I-like enzymes work as $F_{420}H_2$ dehydrogenases [6]. In general, 11 subunits (NuoA to D and NuoH to N), suggested to be responsible for charge translocation and quinone binding, are common to the complexes from the three domains of life. Recently the investigation on complex I gained a new enthusiasm due to the large amount of studies performed including the publication of the crystal structures of the membrane domain of the *E. coli* enzyme [3], and of the bacterial and mitochondrial complexes I, from *T. thermophilus* [7] and *Y. lipolytica* [8], respectively. Additionally, a structure of complex I from *B. taurus* heart mitochondria, a close relative of the human enzyme, was determined by single-particle electron cryo-EM [9].

I.2.1 Complex I is constituted by modular units

Complex I is an example of the modular structure observed in several energy transducing enzymes [10-15]. The prototype complex I from prokaryotes may be viewed as a combination of 14 subunits most of which have been identified in other complexes. Figure I.2.1 schematically represents the L-shape structure of bacterial complex I, constituted by a peripheral and a membrane arm [2, 4].

The peripheral arm has a Y shape with a length of 130 Å and is composed of seven subunits (NuoB to G and NuoI). One of the Y arms is formed by the subunits NuoE and NuoF and the other by the C-terminal domain of NuoG. The central part is formed by the N-terminal domain of NuoG and by the subunits NuoC and NuoI. At the end of the

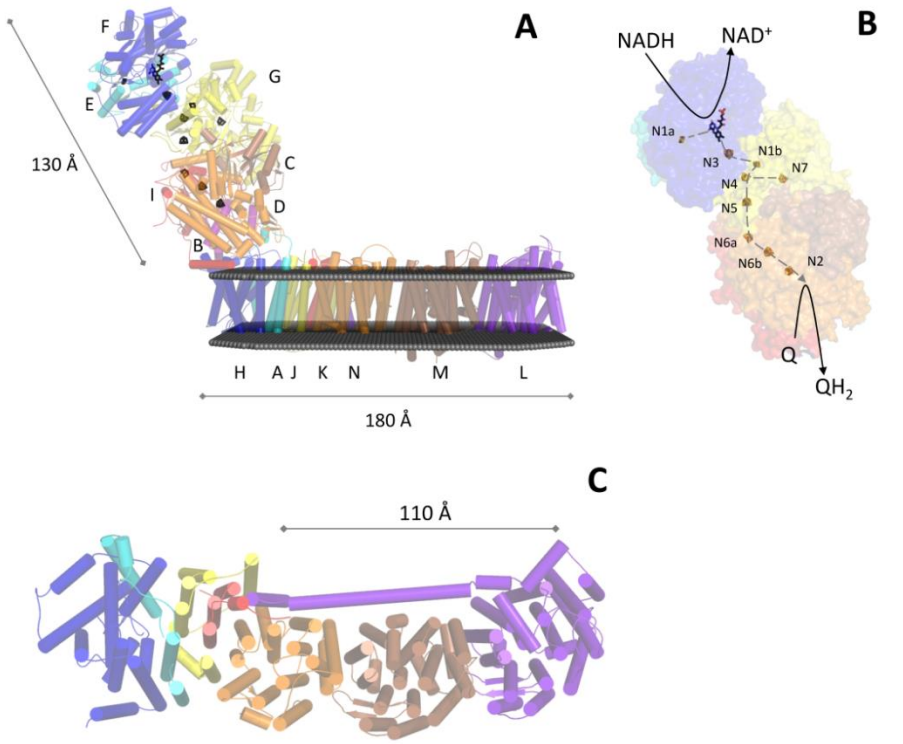


Figure I.2.1: Schematic representation of respiratory complex I. A) Complex I composed of 14 subunits. One FMN and 9 iron-sulfur centers are shown in black. Peripheral arm is composed of: NuoB (purple), NuoC (brown), NuoD (orange), NuoE (cyan), NuoF (blue), NuoG (yellow) and NuoI (red); while the membrane arm is composed of: NuoA (cyan), NuoH (blue), NuoJ (yellow), NuoK (red), NuoL (purple), NuoM (brown) and NuoN (orange). Peripheral arm and membrane arm are 130 Å and 180 Å long. B) Schematic representation of the Complex I peripheral arm. NADH is oxidized by the FMN present in NuoF. The 2 electrons are transferred from the FMN to the quinone binding site (at NuoB and NuoD subunits) through a series of iron-sulfur clusters (N1a to N7). C) Top view of the Complex I membrane arm. A long amphipathic helix is present at the C-terminal of NuoL is 110 Å long. The schematic representation is based on Complex I from *T. thermophilus* (PDB: 4HEA [7]).

Y are the subunits NuoD and NuoB, forming an interface with the membrane domain. NADH is oxidized by the FMN present and the two electrons are transferred to N2 center ([4Fe-4S]) through a series of iron-sulfur clusters, including the binuclear center N1b and the tetranuclear centers N3, N4, N5, N6a and N6b (Figure I.2.1, panel B

and Table I.2.1). The N1a and N7 (may be absent) are not part of the electron pathway wire connecting FMN to N2 center. In addition, except for N1a (-370 mV) and N2 (-50 to -150 mV) centers, all the others centers are isopotential (approximately -250 mV) [2, 16].

Table I.2.1: Prosthetic groups and function of subunits from prokaryotic Complex I.

E. coli	*R. marinus*	*M. mazei*	**Prosthetic groups/ Function**	
NuoA	Nqo7	FpoA	QBS	M
NuoB	Nqo6	FpoB	4Fe-4S (N2) / QBS	P
NuoC[a]	Nqo5	FpoC	unknown	P
NuoD[a]	Nqo4	FpoD	QBS	P
NuoE	Nqo2	-	2Fe-2S (N1a)	P
NuoF	Nqo1	-	FMN + 4Fe4S (N3) / NADH oxidation	P
NuoG	Nqo3	-	2Fe-2S (N1b) + 3x 4Fe-4S (N4, N5, N7[b])	P
NuoH	Nqo8	FpoH	QBS + putative ion transport	M
NuoI	Nqo9	FpoI	2x 4Fe-4S (N6a, N6b)	P
NuoJ	Nqo10	FpoJ	putative ion transport	M
NuoK	Nqo11	FpoK	putative ion transport	M
NuoL	Nqo12	FpoL	ion transport	M
NuoM	Nqo13	FpoM	ion transport	M
NuoN	Nqo14	FpoN	ion transport	M
-	-	FpoO	putative Fe2-S2 binding site	P
-	-	FpoF	$F_{420}H_2$ oxidation	P

[a] The NuoC and NuoD proteins are fused; [b] Do not belong to the electron wire from FMN to the last N2 center and may be not conserved in all Complex I family; QBS) quinone binding site; M) Mebrane; P) Peripheral.

The membrane arm has a curved shape with a total length of 180 Å and is also constituted by seven subunits (NuoA, H, J to N). NuoH is located at the basis of the peripheral arm and the membrane arm is then extended in an almost linear arrangement by subunits

NuoA, NuoJ, NuoK, NuoN, NuoM and NuoL (Figure I.2.1) [4, 17]; whereas the subunits NuoJ to N are homologous to subunits of Mrp Na^+/H^+ antiporters (see I.2.3) [11, 18]. The three homologous antiporter-like subunits, NuoN, M and L, are structurally similar and superimposable [3, 7-9]. The antiporter-like subunit distal to the base of the peripheral arm is NuoL (Figure I.2.1). It has two additional transmembrane helices at the C-terminus which are linked by a ~110 Å long amphipathic α-helix (HL) (Figure I.2.1, panel C) [3]. The role of HL is one of the most discussed topics in the current study of complex I. HL was suggested to have a stability role and/or a functional role in the energy coupling mechanism of complex I [4]. In the latter case, it would work like a piston in order to allow transmission between the base of the peripheral arm, where the quinone is reduced, and the membrane antiporter-like subunits, in which ion translocation is suggested to occur [3] (see I.2.5). Subunits NuoE, F and G constitute the electron input modules and to the exception of the C-terminal domain of NuoG, are homologous to the subunits of soluble NAD^+ reducing hydrogenases. The C-terminal domain of NuoG is homologous to molybdopterin containing enzymes, such as formate dehydrogenase [2]. NuoB, C, D, I and H are homologous to subunits of the membrane-bound [NiFe] hydrogenases (see I.2.2). Several organisms have genes coding for all subunits of Complex I with the exception of those composing the input module (NuoEFG). In some Archaea, for those complexes with F_{420}:phenazine oxidoreductase activity, a flavoprotein (FpoF) (Table I.2.1) has been shown to constitute the input module [6], while for those photosynthetic

complexes with Ferredoxin:plastoquinone oxidoreductase activity, a subunit containing ferredoxin docking site domain (NdhS) has been proposed to constitute the input module [19]. In the case of several other prokaryotic complex I-like enzymes, bacilli and other archaea, the input module is still unknown [20]

I.2.2 Relationship between Complex I and group 4 membrane-bound [NiFe] hydrogenases

A close relationship between the so-called group 4 hydrogenases and complex I has long been recognized [12, 13, 21-24]. Group 4 hydrogenases are membrane-bound enzymes, which receive electrons from cytoplasmatic donors and reduce protons to hydrogen, thus being named hydrogen evolving hydrogenases [23]. The simplest known functional group 4 hydrogenases are the energy-converting hydrogenases (Ech), such as those purified from *Methanosarcina barkeri* and *Thermoanaerobacter tengcongensis* [25-27]. They are constituted by six subunits EchA to F (Figure I.2.2-A). The membrane subunits EchA and EchB are homologous to NuoL and NuoH respectively, while the peripheral subunits EchC, D, E and F (also oriented toward the cytoplasm as in complex I) are homologous to subunits NuoB, C, D and I, respectively (Table I.2.2). The catalytic center is located in subunit EchE and is constituted by a [NiFe] center, which is bound to the protein by specific amino acid residues composing two CxxC motifs, one located close to the N-terminus and the other at the C-terminus. Ech receives electrons from a cytoplasmatic ferredoxin [21].

Formate hydrogen lyase 1 (FHL-1) from *E. coli* includes a group 4 hydrogenase, with the same composition as Ech, that receives electrons from formate *via* a formate dehydrogenase subunit (Fdh-F) [21] (Figure I.2.2-B). Interestingly, the latter protein is homologous to the C-terminus of subunit NuoG from complex I (Table I.2.2). In addition, a gene coding for a protein homologous to the N-terminus of NuoG, and thus possibly involved in electron transfer, is also present in the gene cluster (*hyc*) coding for this hydrogenase [13].

E. coli contains another group 4 hydrogenase (Hyf) encoded by the *hyf* gene cluster (Figure I.2.2-C). This gene cluster is similar to *hyc*, but includes three additional genes [28]. Two of these encode homologues of the antiporter-like subunits from complex I and another (*hyfE*) encodes a protein, whose C-terminus is homologous to NuoK [13]. The electron donor to this hydrogenase is still unknown.

Rhodospirillum rubrum and *Carboxydothermus hydrogenoformans* contain a group 4 hydrogenase, encoded by genes *CooMKLXUH*, whose electron donor is a ferredoxin, reduced by a CO dehydrogenase [26]. These hydrogenases were suggested to conserve energy, allowing *R. rubrum* to grow on CO as the sole energy source [29].

More complex membrane-bound [NiFe] hydrogenases were also identified, such as the membrane-bound hydrogenase (Mbh) from *Pyrococcus furiosus* [30] (Figure I.2.2-D) and the energy-converting hydrogenase A (Eha) from *Methanobacterium thermoautotrophicum* [31].

These hydrogenases are predicted to be composed of 13 or more subunits, having 4 to 6 subunits homologous to complex I [30, 31].

Table I.2.2: Subunits homology between members of Complex I family.

Complex I	Type 4 membrane-bound hydrogenases					Ehr		
E. coli	*M. Mazei*	*E. coli*	*E. coli*	*M. thermoautotrophicum*	*P. furiosus*	*P. furiosus*	*G. sulfurreducens*	
NuoA	-	-	-	-	-	-	-	M
NuoB	EchC	HycG	HyfI	EhaN	MbhJ	MbxJ	EhrS	P
NuoC[a]	EchD	HycE[a]	HyfG[a]	-	MbhK	MbxK	-	P
NuoD[a]	EchE	HycE[a]	HyfG[a]	EhaO	MbhL	MbxL	EhrL	P
NuoE	-	-	-	-	-	-	-	P
NuoF	-	-	-	-	-	-	-	P
NuoG	-	HycB[b]	HyfA[b]	-	-	-	-	P
NuoH	EchB	HycD	HyfC	EhaJ	MbhM	MbxM	EhrB	M
NuoI	EchF	HycF	HyfH	-	MbhN	MbxN	-	P
NuoJ	-	-	-	-	-	-	-	M
NuoK	-	-	HyfE[c]	-	MbhG[c]	MbxG[c]	EhrC	M
NuoL	EchA	HycC	HyfB[e],HyfD[e]	EhaH[d]	MbhH	MbxH[e],MbxH´[e]	EhrA	M
NuoM	-	-	HyfF	-	-	-	EhrD	M
NuoN	-	-	-	-	-	-	-	M
-	-	-	-	-	-	-	-	P
-	-	-	-	-	-	-	-	P
-	-	-	-	-	MbhA	MbxA	-	M
-	-	-	-	-	MbhB	MbxB	-	M
-	-	-	-	-	MbhC	MbxC	-	M
-	-	-	-	-	MbhD[f],MbhE[f], MbhF[f]	MbxD[f],MbxF[f]	-	M
-	-	-	-	EhaA	-	-	-	M
-	-	-	-	EhaB	-	-	-	M
-	-	-	-	EhaC	-	-	-	M
-	-	-	-	EhaD	-	-	-	M
-	-	-	-	EhaE	-	-	-	M
-	-	-	-	EhaF	-	-	-	M
-	-	-	-	EhaG	-	-	-	M
-	-	-	-	EhaI	-	-	-	M
-	-	-	-	EhaK	-	-	-	M
-	-	-	-	EhaL	-	-	-	M
-	-	-	-	EhaM	-	-	-	P
-	-	-	-	EhaP	-	-	-	P
-	-	-	-	-	MbhI	-	-	M

[a] The proteins NuoC and NuoD are fused, [b] This protein is homologous to the N-terminus of NuoG, [c] This protein is homologous to the C-terminus of NuoK, [d] EhaH has only 6 transmembrane helices, [e] These proteins individually are homologous to NuoL, [f] These proteins together are homologous to MrpB. P) peripheral subunit, M) membrane subunit.

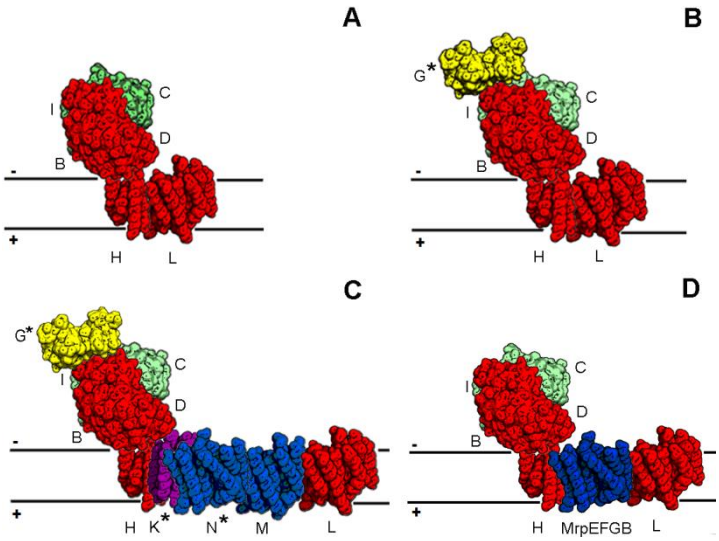


Figure I.2.2: Schematic representation of group 4 membrane bound [NiFe] hydrogenases. A) Energy-converting hydrogenase (Ech). B) Membrane bound hydrogenase Hyc. C) Membrane bound hydrogenase Hyf. D) Membrane bound hydrogenase Mbh. The colour code is: red – homologous subunits to NuoB, D, H, and L (Ech: EchC, EchE, EchB and EchA; Hyc: HycG, HycE, HycD and HycC; Hyf: HyfI, HyfG, HyfC and HyfB; Mbh: MbhJ, MbhL, MbhM and MbhH) (see section II.1); green - homologous subunits to NuoC and NuoI (Ech: EchD and EchF; Hyc: HycE and HycF; Hyf: HyfG and HyfH; Mbh: MbhK and MbhN); blue - homologous subunits to NuoM/N and MrpE, F, G, B (Hyf: HyfD and HyfF; Mbh: MbhA to MbhF); purple - homologous subunit to NuoK (Hyf: HyfE; Mbh: MbhG) and yellow - homologous subunit to NuoG (HycB and HyfA). (G*: Protein is homologous to the C-terminus of NuoG; K*: Protein is homologous to the C-terminus of NuoK; N*: Protein is homologous to antiporter-like). The schematic representation of these complexes is based on the crystal structure of Complex I from *T. thermophilus* (PDB: 4HEA [7]). Table I.2.2.

The presence of group 4 hydrogenases has been reported for several other prokaryotic organisms [23, 32]. Importantly, experiments with inverted membrane vesicles showed that the hydrogenases from *M. mazei* (Ech) and *P. furiosus* (Mbh) coupled H_2 production to the establishment of a transmembrane difference of electrochemical

potential [30, 33], *i.e.* the results showed that these enzymes are capable of energy transduction.

We have recently made a thorough investigation of complex I and group 4 [NiFe] hydrogenases and established unequivocally a third member of this family of proteins, the energy-converting hydrogenase related complexes, Ehrs [34] (see section II.1). These complexes are similar to the hydrogenases but lack the [NiFe] binding site [34, 35] (Table I.2.1).

I.2.3 Relationship between Complex I and Mrp Na^+/H^+ antiporters

Multiple resistance to pH (Mrp) Na^+/H^+ antiporters are considered as secondary active transporters energized by the transmembrane difference of electrochemical potential [36, 37], for which biochemical and biophysical characterization at the protein level is still poor.

This type of antiporters, generally multimeric complexes, was first identified in *Bacillus halodurans* C-125 where it is crucial for cytoplasmatic pH homeostasis in this alkaliphilic bacterium [38]. The *mrp* gene cluster has typically seven genes encoding different integral membrane proteins, named MrpA to G, that were suggested to be coordinately expressed as an operon [39-42]. This entity is called group 1 *mrp* gene cluster [37]. Group 2 *mrp* gene clusters are similar to group 1 but with *mrpB* fused to *mrpA*. Group 3 is more heterogenous, the gene clusters do not have *mrpA*, but may contain multiple copies of *mrpD* [36, 37]. The products of *mrpA* and *mrpD* are

homologous to subunits NuoL/M/N of complex I and that of *mrpC* has homology to NuoK [11, 43]. Surprisingly, the C-terminal domain of MrpA has structural homology to the long amphipathic α-helix and to NuoJ subunit from Complex I [4, 18].

MrpA and MrpD have been shown to have a role in Na^+ resistance and Na^+-dependent pH homeostasis in *Bacillus subtilis* [39]. The functional similarity of MrpA to NuoL and that of MrpD to NuoN was recently corroborated by complementation studies using *B. subtilis* MrpA and MrpD deletions strains, in which the subunits of complex I could replace those of the Mrp antiporter [44, 45]. Recently the Mrp complex from *Bacillus pseudofirmus* OF4 was purified, reconstituted in liposomes and Na^+/H^+ antiporter activity was confirmed [46].

Evolutionary scenarios for complex I have been suggested over the past years, most of which proposing that it resulted from the assembling of a soluble NAD^+ reducing hydrogenase with an Mrp-like antiporter [11, 20, 24].

I.2.4 Ion translocation sites in Complex I

Based on the structural information [3, 7-9] and taking into account the accepted pumping stoichiometry of $4H^+/2e$ [47-49], the presence of four ion translocation sites in Complex I was suggested. Three of these translocation sites were hypothesized to be present in each antiporter-like subunit (NuoL, M and N) and the fourth site was

proposed to be formed by a consortium composed by subunits NuoH, J and K (Figure I.2.3) [7].

Apart from the C-terminus of NuoL, the structures of those three subunits (NuoN, M and L) are superimposable, showing 14 conserved helices containing two inverted repeats in a face-to-back arrangement formed by transmembrane (TM) helices 4 to 8 and 9 to 13 (Figure I.2.3). Each repeat has a discontinuous helix (TM7 and TM12) where a lysine or a glutamate residue is present and a partly unwound helix (TM8) where a lysine or a histidine residue is present (Figure I.2.3). These residues were suggested to be important for charge translocation [3]. Further analysis indicated that an additional ion pathway may exist in the interface between subunits through a conserved glutamate residue in TM5.

NuoH subunit shows eight conserved TM helices and an additional non conserved TM helix at the C-terminal. On contrary to the antiporter-like subunits, NuoH has very small and very long TM helices and with several helices tilted relative to the membrane plane [7]. NuoH has a five TM repeat (TM2 to 6) similar to the ones of the antiporter-like subunits (Figure I.2.3). This repeat forms a half-channel turned to the cytoplasm as the one present in the first repeat of the antiporter-like subunits. Several glutamate residues were found in similar positions to key residues from the antiporter-like subunits, such as, one to the glutamate residue from TM5 and two to the lysine residue present in TM7 [7, 34]. The unique NuoH fold shows an unusual network of negatively charged residues connecting the quinone binding site to NuoAJK subunits. Due to the abundance of

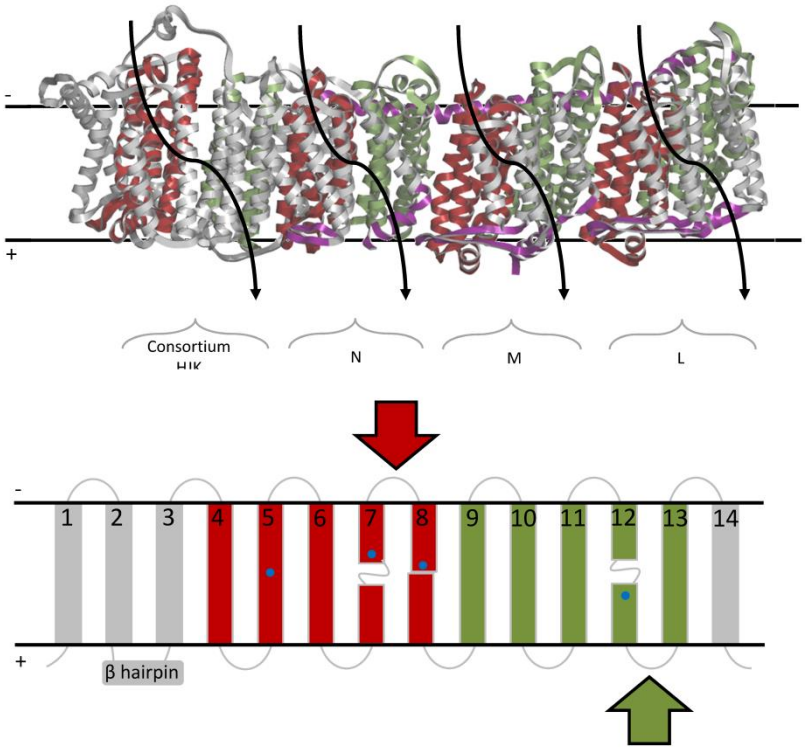


Figure I.2.3: Schematic representation of the four proton pathways in complex I [7]. Top) membrane part of Complex I from *T. thermophilus* and the proposed proton pathways, which are the consortium NuoHJK and 3 antiporter-like subunits (NuoN, M and L); Bottom) First and second half channels (TM4 to 8 and TM9 to 13) of an antiporter-like subunit. The first half channel is facing the cytoplasm, while the second is facing the periplasm. Colored in: Red – 1st half channel; Green – 2nd half channel; purple – proposed structural/connecting elements: HL, β-hairpins and TM14; blue – proposed residues involved in charge translocation. The schematic representation of membrane subunits from Complex I is based on the crystal structure of Complex I from *T. thermophilus* (PDB: 4HEA [7]).

the negatively charged residues, namely glutamate residues, this network was referred as E-channel [7]. The putative fourth ion pathway is formed by NuoH cytoplasmic half-channel (Figure I.2.3, in red) and NuoJ and K periplasmic half-channel, where TM3 from NuoJ corresponds to TM8 from antiporter-like subunits (Figure I.2.3, in green) [4, 7].

The presence of additional elements connecting the antiporter-like subunits was shown by the crystal structure of complex I [3, 7-9].

HL is spanning laterally to the three antiporter-like subunits and ends in a TM helix (TM16) anchored at the interface of NuoN (TM6 and TM7 helices), NuoK (TM1 helix and C-terminal) and NuoJ (TM2 helix) (Figure I.2.3, in purple) [4]. Three β-hairpins are present between TM helices 2 and 3 and localized close to the opposite side of the membrane to that of HL. The β hairpin from one subunit contacts its neighbor by a stretch of a C-terminal amphipathic helix (Figure I.2.3, in purple). These structural arrangements were suggested to contribute to the stability of the complex [3].

I.2.5 Quinone binding site in Complex I

No bound quinone was observed in all obtained crystals from Complex I [3, 7-9]. The quinone binding site is localized in an enclosed cavity formed by the soluble NuoB and D subunits and membrane NuoA and H subunits. Complex I crystal structure from *T. thermophilus* soaked with decylubiquinone [7] shows a long enclosed cavity to bind quinone, which is consistent with previous results [50-53]. Decylubiquinone headgroup binds at ~12 Å from the the N2 iron-sulfur center from NuoB [7]. Corroborating previous labelling experiments [51], the quinone aromatic ring is positioned between conserved tyrosine and histidine residues from NuoD subunit [7].

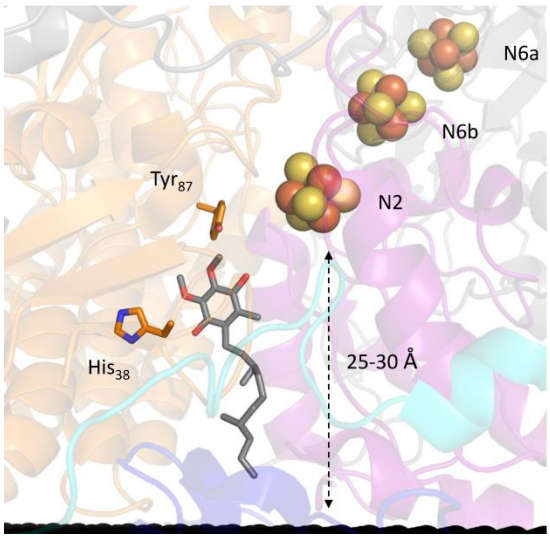


Figure I.2.4: Schematic representation of quinone binding in complex I [7]. View from the distal Na^+/H^+ antiporter-like subunits. Quinone binding cavity is enclosed by NuoA (cyan), NuoB (purple), NuoD (orange), NuoH (blue) subunits of Complex I. Conserved histidine (His38) and tyrosine (Tyr87) residues positions are shown. Spheres) N2, N6a and N6b iron-sulfur centers; quinone (gray) was introduced manually since the Complex I crystallographic structure with the co-crystalized quinone coordination file was not released. The schematic representation of the quinone binding site is based on the crystal structure of Complex I from *T. thermophilus* (PDB: 4HEA [7]).

I.2.6 Ion/electron coupling mechanism in Complex I

In complex I, the coupling process of electron transfer and ion translocation is still poorly understood. Over the past decades different mechanisms have been proposed, several of which may now be excluded based on the available structural data on complex I, namely those suggesting the direct involvements of FMN [54] or of the iron-sulfur center N2 [55]. In addition, models describing H^+ translocation directly linked to the quinone/quinol couple are no longer consensually considered [56]. The recently available structural data do not support “strictly direct” coupling models. Alternatively, a

combination of a direct mechanism, involving a quinone or N2 center, and an indirect mechanism, with the intervention of the antiporter-like subunits, has emerged. However, the involvement of the quinone or of the N2 center is unlikely by the observation of the location of N2 center (~25-30 Å away from the membrane surface) and of the quinone binding site ~15 Å away from the membrane surface [3, 4, 53].

Strictly indirect mechanisms are presently the most attractive operating mode of complex I. This type of mechanism is supported by the structural data, conformational changes monitored upon NADH binding and/or quinone reduction [57, 58], and the observation that the catalytic and transport activities can be decoupled [59] (see section II.3). The crystallographic structures of complex I show the presence of a HL parallel to the membrane, which is suggested to act as a transmission element connecting possible conformational changes near the quinone reduction site to the antiporter-like subunits (Figure I.2.3). Efremov *et al* proposed that a piston-like motion of the HL results in H^+ translocation [4]. Afterwards, Brandt proposed a model in which the H^+ are translocated by two different pump modules, being these connected by the HL in a similar way to a rod of a steam engine [60]. Both models were challenged by studies using site directed mutants, which suggested that the HL, instead of being a transmission element, may be just needed to keep the subunits together during catalytic turnover [61-63]. However, *E. coli* complex I mutated in an aspartate residue of the HL, which is close to NuoM subunit, presented approximately half of the H^+/e stoichiometry, which was interpreted as the transmission between HL

and NuoL and NuoM being hampered [61]. Particularly, the HL was proposed to work as a long helical energy transmission element connecting the antiporter-like subunits (NuoL, M and N) [63]. Based on the observations that complex I from *Y. lipolytica,* missing two of the antiporter-like subunits (and the HL), was still catalytically active and capable of proton translocation, although at half the stoichiometry of the complete enzyme, Drose *et al* suggested that complex I has two distinct ion pump modules connected in series by the HL, which would work as a transmission component [63].

In addition, E-channel in NuoH composed of a series of negative charged residues, is connected to conserved charged residues at the middle of the membrane subunits, including NuoA, J, K, N, M and L. This feature suggests that a conformational change at the quinone binding site may be transduced trough these charged residues, including the lysine residues in TM7 from the antiporters, to the distal part of the membrane arm. [7, 64, 65].

Recently, large-scale molecular simulation was performed using the solved crystallographic structures of Complex I from *E. coli* and *T. thermophilus*. They suggest that the initial step of conformational changes involve a charge imbalance arising from the quinone reduction, and that it leads to long range energy propagation through reorientations of charged residues present in the middle of the membrane subunits [66, 67].

Although the structure/function relation in complex I is still not clear, it is (almost) consensual that the coupling process in complex I

occurs through an indirect mechanism. In such a situation the transport of other ions, namely Na^+, beside H^+ cannot be excluded.

I.2.7 H^+ as the coupling ion of Complex I

Complex I is generally accepted to contribute to the transmembrane difference of electrochemical potential by translocating H^+ ions across the membrane, in a process that is coupled to its redox activity. However, the determination of the number of H^+ ions or charges translocated *per* catalytic turnover (corresponding to $2e^-$ transfer between NADH and quinone) has not been a trivial task.

Early observations by Lawford and Garland (1972) and Ragan and Hinkle (1975) indicated a maximal $H^+/2e^-$ stoichiometry of 2 for rat liver and bovine heart mitochondrial complex I [68, 69]. Later, studies performed by the groups of Lehninger [70] and Azzone [71] indicated a stoichiometry of approximately $4H^+/2e^-$ in rat liver and heart mitochondria, which is now generally accepted [47, 48, 72-74]. The same value was obtained for complex I from the yeast *Y. lipolytica*, either using intact mitochondria or purified enzyme reconstituted into liposomes [49]. However, Wikström and Hummer using the values of the H^+/ATP ratio of the $F1F_O$-ATP synthase of animal mitochondria and the measured $ATP/2e^-$ ratios for different segments of the mitochondrial respiratory chain, calculated a value of 3 $H^+/2e^-$ for mitochondrial complex I [75]. Indeed, such stoichiometry was observed by Jonge and Westerhoff [76] using submitochondrial particles. Only two studies addressed the $H^+/2e^-$ stoichiometry by

bacterial complexes I; a stoichiometry of 3 to 4 $H^+/2e^-$ was determined for complex I of aerobically grown *Paracoccus denitrificans* [77], while a value of 3 was indirectly determined for dimethyl-sulfoxide respiring *E. coli* cells [78]. A $H^+/2e^-$ stoichiometry of 3 for complex I does not agree with the existence of 4 translocating sites if all of these are considered to be active.

I.2.8 Na^+ as the coupling ion of Complex I

The established role of H^+ as coupling ion in complex I was challenged by the studies of Steuber and coworkers using two closely related organisms, *K. pneumonia* and *E. coli* [79]. By atomic absorption spectroscopy and radioactive $^{22}Na^+$ techniques, the authors observed NADH-driven Na^+ uptake in inverted membrane vesicles of *K. pneumonia*. This Na^+ transport was abolished by the complex I inhibitor, rotenone, but not in the presence of the protonophore CCCP. These observations were indicative of a primary Na^+ pump activity by complex I and not a combined action of a proton pump and Na^+/H^+ antiporters [80]. Similar conclusions were obtained using a Na^+/H^+ antiporter deficient *E. coli* strain. In this study, the authors observed an increase of Na^+ tolerance of the studied strain when the expression of complex I was higher. Na^+ translocation by complex I was confirmed by using the complex I inhibitor rotenone and a mutant devoid of complex I [79]. Interestingly, it was observed that the C-terminally truncated subunit NuoL of *E. coli* complex I (reconstituted into liposomes [81] or expressed in S*accharomyces cerevisiae* [82]) performed passive Na^+ uptake.

Complex I of *K. pneumonia* was reconstituted into liposomes and a $Na^+/2e^-$ ratio of 2 was determined [83]. When complex I from *K. pneumonia* and the Na^+-dependent ATP synthase from *Ilyobacter tartaricus* were reconstituted together into liposomes, it was observed that the $\Delta\mu Na^+$ generated by NADH-driven Na^+ translocation by complex I could serve as driving force for the ATP synthesis and that the $\Delta\mu Na^+$ generated by ATP hydrolysis could drive NADH formation by reversed electron transfer in complex I [84].

Na^+ pumping activity was also detected in complex I of *Y. lypolitica* and again, NADH-driven Na^+ translocation from the negatively charged side to the positively charged side of the membrane of submitochondrial particles was observed [85].

All these finding were highly debated because in *K. pneumonia* [86], *E. coli* [87] and *Y. lypolitica* [49, 88], H^+ translocation from the negative to the positive side of the membrane was observed. Moreover, Na^+ translocation was attributed to Na^+-translocating NADH:quinone oxidoreductase (Na^+-NQR) in *K. pneumonia* [86] and to secondary Na^+/H^+ antiporter activity in *E. coli* [87]. Nevertheless, growth and experimental conditions were different among the different studies, which might strongly influence the expression of respiratory enzymes and the nature of the coupling ion.

I.2.9 Na^+/H^+ antiporter activity of Complex I

The first suggestion of a Na^+/H^+ antiporter activity by complex I was made by Stolpe and Friedrich using isolated *E. coli* complex I reconstituted into liposomes [87]. The results showed NADH-driven H^+

translocation, which increased in the presence of Na^+. The authors suggested that a putative translocation of Na^+ by complex I drives further the translocation of H^+ in the opposite direction, increasing ΔpH. These results indicated the role of H^+ as the coupling ion of complex I and a possible Na^+/H^+ antiporter activity by complex I in *E. coli*.

NADH-driven Na^+/H^+ antiporter activity of complex I was directly observed using a ^{23}Na-NMR spectroscopy method developed by our group [89-91]. This technique allowed the direct monitoring of the changes in Na^+ concentration inside and outside of inverted membrane vesicles of *R. marinus* and *E. coli* [89-91]. The results showed proton translocation in the same direction of the established ΔΨ and, for the first time, translocation of Na^+ by complex I in the direction opposite to that of the established ΔΨ. In these systems, Na^+ translocation increased in the presence of the protonopohore CCCP, showing that Na^+ transport is not dependent on ΔpH and thus is a primary event. The presence of Na^+ was also observed to increase the ΔpH, although it was not necessary for the catalytic or proton transport activities, which was in agreement with the observations by Stolpe and Friedrich for *E. coli* [87].

The Na^+/H^+ antiporter activity of complex I of *R. marinus* was further corroborated by investigations with the Na^+/H^+ antiporter inhibitor EIPA. This inhibitor promoted the decoupling of the catalytic and transport activities of complex I from *R. marinus* and two different inhibition profiles of H^+ translocation, in the presence and absence of Na^+, were observed [59] (see section II.3). We hypothesized that

complexes I from *E. coli* and *R. marinus* have two types of energy coupling sites, one Na^+ independent, working as proton pump and the other Na^+ dependent, working as Na^+/H^+ antiporter. Recently, the ion translocation activity by complex I from *E. coli* devoid of subunit NuoL was investigated and we observed that the mutated complex I does not translocate Na^+ and is less effective in proton pumping [92] (see section II.4). These studies corroborated the idea that H^+ is the coupling ion in complex I from *E. coli* and that subunit NuoL may be involved in the translocation of H^+ and Na^+ ions, possibly acting as a Na^+/H^+ antiporter or influencing the activity of other subunits of complex I. In addition, these observations are in agreement with those by Friedrich and coworkers, who previously reconstituted into liposomes an *E. coli* ΔNuoL mutant of complex I and obtained half of the value of the $H^+/2e^-$ stoichiometry of wild-type complex I [93].

We also investigated Na^+ translocation by complex I from *P. denitrificans*. In this case we observed translocation of H^+ but not of Na^+. Knowing that the two complexes I for which we observed Na^+ translocation use menaquinone as substrate, and the one from *P. denitrificans* uses ubiquinone, we hypothesized a correlation between the type of quinone used as substrate and the Na^+/H^+ antiporter activity [90]. Hirst and coworkers reconstituted complex I from *B. taurus* heart mitochondria (which reduces ubiquinone) into liposomes and observed H^+ translocation by complex I in its "deactive" form (*i.e.* no redox activity due to lack of substrates), when a Na^+ electrochemical potential was imposed [94]. This suggested that mitochondrial complex I may perform Na^+/H^+ antiporter activity under

certain conditions. The conversion between an "active" and a "deactive" form of complex I is a characteristic of some eukaryotes, never observed in bacteria [95]. Hirst and coworkers suggested that the redox and translocation modules of complex I become disconnected in the "deactive" form, possibly allowing the antiporter-like subunits of complex I to perform their native function (*i.e.* Na^+/H^+ antiporter activity) [94].

I.2.10 Thermodynamic considerations

Since complexes I from *E. coli*, *R. marinus* and *P. denitrificans* are very similar, the differences observed in terms of Na^+ translocation may be related to the metabolic characteristics of each organism, namely the quinones used by complex I. Both *E. coli* (grown under anaerobic or microaerophylic conditions) and *R. marinus* complexes I reduce menaquinone (Em,7≈ −80 mV), whereas complex I from *P. denitrificans* uses ubiquinone (Em,7≈ +100 mV) as substrate [96]. The difference in the reduction potentials of the two types of quinones has strong thermodynamic implications. Considering the reduction potential of the NAD^+/NADH pair (Em,7≈ −320 mV), the redox potential difference (ΔE) between NADH and menaquinone is 240 mV and between NADH and ubiquinone is 420 mV [97, 98] (Figure I.2.5). As can be observed, the energy available for proton translocation is very different in these two situations. Considering that in both cases the $H^+/2e^-$ stoichiometry would be the same, an immediate raised question was whether Δp would be the same in organisms that produce menaquinone and in those that synthesize ubiquinone.

Unfortunately, the information on the determination of Δp for different organisms or growth conditions is scarce. As far as we know, the only study that addresses this question is the work by Tran and Unden [97], which determined that in *E. coli*, Δp is 160 mV and 140 mV in aerobic and anaerobic respiration, respectively. *E. coli* uses ubiquinone in aerobic conditions and menaquinone in anaerobic conditions. Thus, these results do not suggest that Δp of menaquinone reducing organisms would be much different from that of ubiquinone reducing organisms.

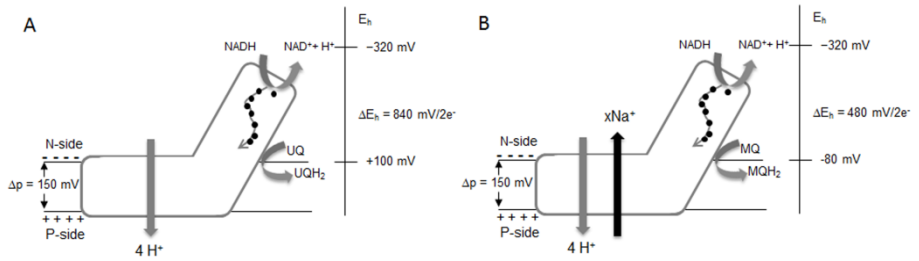


Figure I.2.5: Thermodynamic contributions for complex I using menaquinone (MQ) or ubiquinone (UQ) as substrate. Schematic representation of ubiquinone-reducing (A) and menaquinone-reducing (B) complexes I. The free energy obtained from the redox reaction drives ion translocation across the membrane, against an average Δp of 150 mV. Considering the reduction potential of the NAD+/NADH pair (Em,7≈ −320 mV), the redox potential difference (ΔE) between NADH and menaquinone is 240 mV and between NADH and ubiquinone is 420 mV. For a two electron redox reaction, ΔE is related to Δp by the following expression, nΔp=2ΔE, where n is the number of protons that can be pumped. The menaquinone-reducing complexes I would not have enough available energy to reach the accepted stoichiometry of 4 $H^+/2e^-$. We hypothesize that the dissipation of Na^+ chemical gradient (Δp) would provide the required energy. In ubiquinone-reducing complexes I enough ΔE is already available by the redox reaction and only protons are translocated.

For a two electron redox reaction, ΔE relates to the number of protons (n) that can be translocated and to the Δp by the relation, nΔp=2ΔE [98].

If we consider an average Δp value of 150 mV, the oxidation of menaquinone and ubiquinone allows a maximal $H^+/2e^-$ stoichiometry of 3 and 5, respectively (Figure I.2.5). However, it is important to refer that these calculations were performed based on the midpoint potentials of the substrates, considering that the pools of substrates are 50 % oxidized and 50 % reduced. In physiological conditions, the $NADH/NAD^+$ ratio is smaller than 1 (either in aerobic or anaerobic conditions) and then the actual reduction potential for the $NADH/NAD^+$ couple would be higher, possibly leading to slightly smaller $H^+/2e^-$ stoichiometries than those we present here [99].

In the organisms that use ubiquinone as the electron acceptor (e.g. *P. denitrificans* and eukaryotes), the energy obtained by the reduction of ubiquinone seems to fulfill the stoichiometry of $4H^+/2e^-$. In the case of organisms that use menaquinone (e.g. *E. coli* and *R. marinus*), the energy obtained by the menaquinone reduction is only sufficient for a stoichiometry of $3H^+/2e^-$, which is in accordance with the value suggested by Bogachev *et al* when using DMSO respiring *E. coli* cells [78]. Nevertheless, the crystal structure of complex I from *T. thermophilus* (a menaquinone reducing organism) suggests the presence of four ion channels, which may allow the translocation of 4 protons *per* catalytic turnover [7].

We hypothesized that in the menaquinone reducing complex I, at least one of the coupling sites could function as a Na^+/H^+ antiporter. The energy released by dissipation of Na^+ membrane potential would allow the translocation of additional protons in order to achieve the stoichiometry of $4H^+/2e^-$ (Figure I.2.5). A simple exchange of 1 H^+ per 1

Na^+ would not contribute to an increase of the electrical component of the electrochemical potential but would contribute to increase ΔpH. If the Na^+/H^+ antiporter activity of complex I would be electrogenic *i.e.,* Na^+/H^+ stoichiometry different from 1, the Na^+/H^+ antiporter activity could also contribute to ΔΨ.

Menaquinone reducing complexes are present in early branches of prokaryotic phyla and can be considered ancestor enzymes, relatively to ubiquinone reducing complexes [100]. When the substrate changed from menaquinone to ubiquinone, the NADH-quinone redox reaction provided more energy for ion translocation, and the complex I would have eventually lost the need to use ΔpNa to promote additional H^+ translocation (Figure I.2.5).

Na^+/H^+ antiporters were reported to change its directionality and stoichiometry in response to the environmental conditions. In *E. coli*, an electroneutral (1:1) H^+/Na^+ antiporter operates under acidic and neutral conditions and an electrogenic activity (H^+/Na^+= 2) with export of Na^+ is detected under alkaline conditions. Taking this into consideration, the possibility of Na^+ translocation to the positive side of the membrane may also be put forward. Whether this would contribute to the establishment of a transmembrane difference of electrochemical potential depends on the stoichiometry of the transport [101]. The idea of a reversible Na^+/H^+ antiporter could partially solve the apparent controversy of the nature of the coupling ion.

I.2.10 References

1. Schapira, A. H. (1998) Human complex I defects in neurodegenerative diseases, *Biochim Biophys Acta.* 1364, 261-70.

2. Sazanov, L. A. & Hinchliffe, P. (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, *Science.* 311, 1430-6.

3. Efremov, R. G. & Sazanov, L. A. (2011) Structure of the membrane domain of respiratory complex I, *Nature.* 476, 414-20.

4. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I, *Nature.* 465, 441-5.

5. Kerscher, S., Drose, S., Zickermann, V. & Brandt, U. (2008) The three families of respiratory NADH dehydrogenases, *Results Probl Cell Differ.* 45, 185-222.

6. Baumer, S., Ide, T., Jacobi, C., Johann, A., Gottschalk, G. & Deppenmeier, U. (2000) The F420H2 dehydrogenase from *Methanosarcina mazei* is a Redox-driven proton pump closely related to NADH dehydrogenases, *J Biol Chem.* 275, 17968-73.

7. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. (2013) Crystal structure of the entire respiratory complex I, *Nature.* 494, 443-8.

8. Zickermann, V., Wirth, C., Nasiri, H., Siegmund, K., Schwalbe, H., Hunte, C. & Brandt, U. (2015) Mechanistic insight from the crystal structure of mitochondrial complex I, *Science.* 347, 44-9.

9. Vinothkumar, K. R., Zhu, J. & Hirst, J. (2014) Architecture of mammalian respiratory complex I, *Nature.* 515, 80-4.

10. Refojo, P. N., Sousa, F. L., Teixeira, M. & Pereira, M. M. (2010) The alternative complex III: a different architecture using known building modules, *Biochim Biophys Acta.* 1797, 1869-76.

11. Mathiesen, C. & Hagerhall, C. (2003) The 'antiporter module' of respiratory chain complex I includes the MrpC/NuoK subunit -- a revision of the modular evolution scheme, *FEBS Lett.* 549, 7-13.

12. Friedrich, T. & Weiss, H. (1997) Modular evolution of the respiratory NADH:ubiquinone oxidoreductase and the origin of its modules, *J Theor Biol.* 187, 529-40.

13. Efremov, R. G. & Sazanov, L. A. (2012) The coupling mechanism of respiratory complex I - A structural and evolutionary perspective, *Biochim Biophys Acta*, doi:10.1016/j.bbabio.2012.02.015.

14. Rothery, R. A., Workun, G. J. & Weiner, J. H. (2008) The prokaryotic complex iron-sulfur molybdoenzyme family, *Biochim Biophys Acta.* 1778, 1897-929.

15. Baymann, F., Lebrun, E., Brugna, M., Schoepp-Cothenet, B., Giudici-Orticoni, M. T. & Nitschke, W. (2003) The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes, *Philos Trans R Soc Lond B Biol Sci.* 358, 267-74.

16. Ohnishi, T. (1998) Iron-sulfur clusters/semiquinones in complex I, *Biochim Biophys Acta.* 1364, 186-206.

17. Holt, P. J., Morgan, D. J. & Sazanov, L. A. (2003) The location of NuoL and NuoM subunits in the membrane domain of the Escherichia coli complex I: implications for the mechanism of proton pumping, *J Biol Chem.* 278, 43114-20.

18. Virzintiene, E., Moparthi, V. K., Al-Eryani, Y., Shumbe, L., Gorecki, K. & Hagerhall, C. (2013) Structure and function of the C-terminal domain of MrpA in the *Bacillus subtilis* Mrp-antiporter complex--the evolutionary progenitor of the long horizontal helix in complex I, *FEBS Lett.* 587, 3341-7.

19. He, Z., Zheng, F., Wu, Y., Li, Q., Lv, J., Fu, P. & Mi, H. (2015) NDH-1L interacts with ferredoxin via the subunit NdhS in *Thermosynechococcus elongatus*, *Photosynthesis research.* 126, 341-9.

20. Friedrich, T. & Scheide, D. (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases, *FEBS Lett.* 479, 1-5.

21. Hedderich, R. & Forzi, L. (2005) Energy-converting [NiFe] hydrogenases: more than just H2 activation, *J Mol Microbiol Biotechnol.* 10, 92-104.

22. Hedderich, R. (2004) Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I, *J Bioenerg Biomembr.* 36, 65-75.

23. Vignais, P. M. & Billoud, B. (2007) Occurrence, classification, and biological function of hydrogenases: an overview, *Chem Rev.* 107, 4206-72.

24. Moparthi, V. K. & Hagerhall, C. (2011) The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits, *J Mol Evol.* 72, 484-97.

25. Meuer, J., Bartoschek, S., Koch, J., Kunkel, A. & Hedderich, R. (1999) Purification and catalytic properties of Ech hydrogenase from *Methanosarcina barkeri*, *Eur J Biochem.* 265, 325-35.

26. Soboh, B., Linder, D. & Hedderich, R. (2004) A multisubunit membrane-bound [NiFe] hydrogenase and an NADH-dependent Fe-only hydrogenase in the fermenting bacterium *Thermoanaerobacter tengcongensis*, *Microbiology.* 150, 2451-63.

27. Kunkel, A., Vorholt, J. A., Thauer, R. K. & Hedderich, R. (1998) An *Escherichia coli* hydrogenase-3-type hydrogenase in methanogenic archaea, *Eur J Biochem.* 252, 467-76.

28. Andrews, S. C., Berks, B. C., McClay, J., Ambler, A., Quail, M. A., Golby, P. & Guest, J. R. (1997) A 12-cistron *Escherichia coli* operon (hyf) encoding a putative proton-translocating formate hydrogenlyase system, *Microbiology.* 143 (Pt 11), 3633-47.

29. Kerby, R. L., Ludden, P. W. & Roberts, G. P. (1995) Carbon monoxide-dependent growth of *Rhodospirillum rubrum*, *J Bacteriol.* 177, 2241-4.

30. Sapra, R., Bagramyan, K. & Adams, M. W. (2003) A simple energy-conserving system: proton reduction coupled to proton translocation, *Proc Natl Acad Sci U S A.* 100, 7545-50.

31. Tersteegen, A. & Hedderich, R. (1999) *Methanobacterium thermoautotrophicum* encodes two multisubunit membrane-bound [NiFe] hydrogenases. Transcription of the operons and sequence analysis of the deduced proteins, *Eur J Biochem.* 264, 930-43.

32. Pereira, I. A., Ramos, A. R., Grein, F., Marques, M. C., da Silva, S. M. & Venceslau, S. S. (2011) A comparative genomic analysis of energy metabolism in sulfate reducing bacteria and archaea, *Front Microbiol.* 2, 69.

33. Welte, C., Kratzer, C. & Deppenmeier, U. (2010) Involvement of Ech hydrogenase in energy conservation of *Methanosarcina mazei*, *FEBS J.* 277, 3396-403.

34. Marreiros, B. C., Batista, A. P., Duarte, A. M. & Pereira, M. M. (2012) A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases, *Biochim Biophys Acta*.

35. Coppi, M. V. (2005) The hydrogenases of *Geobacter sulfurreducens*: a comparative genomic perspective, *Microbiology.* 151, 1239-54.

36. Krulwich, T. A., Hicks, D. B. & Ito, M. (2009) Cation/proton antiporter complements of bacteria: why so large and diverse?, *Mol Microbiol.* 74, 257-60.

37. Swartz, T. H., Ikewada, S., Ishikawa, O., Ito, M. & Krulwich, T. A. (2005) The Mrp system: a giant among monovalent cation/proton antiporters?, *Extremophiles.* 9, 345-54.

38. Hamamoto, T., Hashimoto, M., Hino, M., Kitada, M., Seto, Y., Kudo, T. & Horikoshi, K. (1994) Characterization of a gene responsible for the Na+/H+ antiporter system of alkalophilic Bacillus species strain C-125, *Mol Microbiol.* 14, 939-46.

39. Ito, M., Guffanti, A. A., Oudega, B. & Krulwich, T. A. (1999) mrp, a multigene, multifunctional locus in *Bacillus subtilis* with roles in resistance to cholate and to Na+ and in pH homeostasis, *J Bacteriol.* 181, 2394-402.

40. Hiramatsu, T., Kodama, K., Kuroda, T., Mizushima, T. & Tsuchiya, T. (1998) A putative multisubunit Na+/H+ antiporter from *Staphylococcus aureus*, *J Bacteriol.* 180, 6642-8.

41. Morino, M., Natsui, S., Swartz, T. H., Krulwich, T. A. & Ito, M. (2008) Single gene deletions of mrpA to mrpG and mrpE point mutations affect activity of the Mrp Na+/H+ antiporter of alkaliphilic Bacillus and formation of hetero-oligomeric Mrp complexes, *J Bacteriol.* 190, 4162-72.

42. Kajiyama, Y., Otagiri, M., Sekiguchi, J., Kosono, S. & Kudo, T. (2007) Complex formation by the mrpABCDEFG gene products, which constitute a principal Na+/H+ antiporter in *Bacillus subtilis*, *J Bacteriol.* 189, 7511-4.

43. Mathiesen, C. & Hagerhall, C. (2002) Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters, *Biochim Biophys Acta.* 1556, 121-32.

44. Moparthi, V. K., Kumar, B., Mathiesen, C. & Hagerhall, C. (2011) Homologous protein subunits from Escherichia coli NADH:quinone oxidoreductase can functionally replace MrpA and MrpD in *Bacillus subtilis*, *Biochim Biophys Acta.* 1807, 427-36.

45. Moparthi, V. K., Kumar, B., Al-Eryani, Y., Sperling, E., Gorecki, K., Drakenberg, T. & Hagerhall, C. (2014) Functional role of the MrpA- and MrpD-homologous protein subunits in enzyme complexes evolutionary related to respiratory chain complex I, *Biochim Biophys Acta.* 1837, 178-85.

46. Morino, M., Suzuki, T., Ito, M. & Krulwich, T. A. (2014) Purification and functional reconstitution of a seven-subunit mrp-type na+/h+ antiporter, *J Bacteriol.* 196, 28-35.

47. Wikstrom, M. (1984) Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone, *FEBS Lett.* 169, 300-4.

48. Galkin, A. S., Grivennikova, V. G. & Vinogradov, A. D. (1999) H+/2e- stoichiometry in NADH-quinone reductase reactions catalyzed by bovine heart submitochondrial particles, *FEBS Lett.* 451, 157-61.

49. Galkin, A., Drose, S. & Brandt, U. (2006) The proton pumping stoichiometry of purified mitochondrial complex I reconstituted into proteoliposomes, *Biochim Biophys Acta.* 1757, 1575-81.

50. Yano, T., Dunham, W. R. & Ohnishi, T. (2005) Characterization of the delta muH+-sensitive ubisemiquinone species (SQ(Nf)) and the interaction with cluster N2: new insight into the energy-coupled electron transfer in complex I, *Biochemistry.* 44, 1744-54.

51. Tocilescu, M. A., Fendel, U., Zwicker, K., Drose, S., Kerscher, S. & Brandt, U. (2010) The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction, *Biochim Biophys Acta.* 1797, 625-32.

52. Zickermann, V., Bostina, M., Hunte, C., Ruiz, T., Radermacher, M. & Brandt, U. (2003) Functional implications from an unexpected position of the 49-kDa subunit of NADH:ubiquinone oxidoreductase, *J Biol Chem.* 278, 29072-8.

53. Hunte, C., Zickermann, V. & Brandt, U. (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I, *Science.* 329, 448-51.

54. Vinogradov, A. D. (2001) Respiratory complex I: structure, redox components, and possible mechanisms of energy transduction, *Biochemistry (Mosc).* 66, 1086-97.

55. Brandt, U. (1997) Proton-translocation by membrane-bound NADH:ubiquinone-oxidoreductase (complex I) through redox-gated ligand conduction, *Biochim Biophys Acta.* 1318, 79-91.

56. Dutton, P. L., Moser, C. C., Sled, V. D., Daldal, F. & Ohnishi, T. (1998) A reductant-induced oxidation mechanism for complex I, *Biochim Biophys Acta.* 1364, 245-57.

57. Pohl, T., Schneider, D., Hielscher, R., Stolpe, S., Dorner, K., Kohlstadt, M., Bottcher, B., Hellwig, P. & Friedrich, T. (2008) Nucleotide-induced conformational changes in the *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I), *Biochem Soc Trans.* 36, 971-5.

58. Mamedova, A. A., Holt, P. J., Carroll, J. & Sazanov, L. A. (2004) Substrate-induced conformational change in bacterial complex I, *J Biol Chem.* 279, 23830-6.

59. Batista, A. P., Marreiros, B. C. & Pereira, M. M. (2011) Decoupling of the catalytic and transport activities of complex I from *Rhodothermus marinus* by sodium/proton antiporter inhibitor, *ACS Chem Biol.* 6, 477-83.

60. Brandt, U. (2011) A two-state stabilization-change mechanism for proton-pumping complex I, *Biochim Biophys Acta.* 1807, 1364-9.

61. Steimle, S., Willistein, M., Hegger, P., Janoschke, M., Erhardt, H. & Friedrich, T. (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I, *FEBS Lett.* 586, 699-704.

62. Belevich, G., Knuuti, J., Verkhovsky, M. I., Wikstrom, M. & Verkhovskaya, M. (2011) Probing the mechanistic role of the long alpha-helix in subunit L of respiratory Complex I from *Escherichia coli* by site-directed mutagenesis, *Mol Microbiol.* 82, 1086-95.

63. Drose, S., Krack, S., Sokolova, L., Zwicker, K., Barth, H. D., Morgner, N., Heide, H., Steger, M., Nubel, E., Zickermann, V., Kerscher, S., Brutschy, B., Radermacher, M. & Brandt, U. (2011) Functional dissection of the proton pumping modules of mitochondrial complex I, *PLoS biology.* 9, e1001128.

64. Euro, L., Belevich, G., Verkhovsky, M. I., Wikstrom, M. & Verkhovskaya, M. (2008) Conserved lysine residues of the membrane subunit NuoM are involved in

energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (Complex I), *Biochim Biophys Acta.* 1777, 1166-72.

65. Verkhovskaya, M. & Bloch, D. A. (2013) Energy-converting respiratory Complex I: on the way to the molecular mechanism of the proton pump, *The international journal of biochemistry & cell biology.* 45, 491-511.

66. Sharma, V., Belevich, G., Gamiz-Hernandez, A. P., Rog, T., Vattulainen, I., Verkhovskaya, M. L., Wikstrom, M., Hummer, G. & Kaila, V. R. (2015) Redox-induced activation of the proton pump in the respiratory complex I, *Proc Natl Acad Sci U S A.* 112, 11571-6.

67. Kaila, V. R., Wikstrom, M. & Hummer, G. (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I, *Proc Natl Acad Sci U S A.* 111, 6988-93.

68. Lawford, H. G. & Garland, P. B. (1972) Proton translocation coupled to quinone reduction by reduced nicotinamide--adenine dinucleotide in rat liver and ox heart mitochondria, *Biochem J.* 130, 1029-44.

69. Ragan, C. I. & Hinkle, P. C. (1975) Ion transport and respiratory control in vesicles formed from reduced nicotinamide adenine dinucleotide coenzyme Q reductase and phospholipids, *J Biol Chem.* 250, 8472-6.

70. Reynafarje, B. & Lehninger, A. L. (1978) The K+/site and H+/site stoichiometry of mitochondrial electron transport, *J Biol Chem.* 253, 6331-4.

71. Pozzan, T., Miconi, V., Di Virgilio, F. & Azzone, G. F. (1979) H+/site, charge/site, and ATP/site ratios at coupling sites I and II in mitochondrial e- transport, *J Biol Chem.* 254, 12000-5.

72. Galkin, A. S., Grivennikova, V. G. & Vinogradov, A. D. (2001) H+/2e- stoichiometry of the nadh:ubiquinone reductase reaction catalyzed by submitochondrial particles, *Biochemistry (Mosc).* 66, 435-43.

73. Brown, G. C. & Brand, M. D. (1988) Proton/electron stoichiometry of mitochondrial complex I estimated from the equilibrium thermodynamic force ratio, *Biochem J.* 252, 473-9.

74. Ripple, M. O., Kim, N. & Springett, R. (2013) Mammalian complex I pumps 4 protons per 2 electrons at high and physiological proton motive force in living cells, *J Biol Chem.* 288, 5374-80.

75. Wikstrom, M. & Hummer, G. (2012) Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications, *Proc Natl Acad Sci U S A.* 109, 4431-6.

76. de Jonge, P. C. & Westerhoff, H. V. (1982) *THe proton-per-electron stoicheiometry of 'site 1' of oxidative phosphorylation at high protonmotive force is close to 1.5*.

77. Meijer, E. M., van Verseveld, H. W., van der Beek, E. G. & Stouthamer, A. H. (1977) Energy conservation during aerobic growth in *Paracoccus denitrificans*, *Arch Microbiol.* 112, 25-34.

78. Bogachev, A. V., Murtazina, R. A. & Skulachev, V. P. (1996) H+/e- stoichiometry for NADH dehydrogenase I and dimethyl sulfoxide reductase in anaerobically grown *Escherichia coli* cells, *J Bacteriol.* 178, 6233-7.

79. Steuber, J., Schmid, C., Rufibach, M. & Dimroth, P. (2000) Na+ translocation by complex I (NADH:quinone oxidoreductase) of *Escherichia coli*, *Mol Microbiol.* 35, 428-34.

80. Krebs, W., Steuber, J., Gemperli, A. C. & Dimroth, P. (1999) Na+ translocation by the NADH:ubiquinone oxidoreductase (complex I) from *Klebsiella pneumoniae*, *Mol Microbiol.* 33, 590-8.

81. Steuber, J. (2003) The C-terminally truncated NuoL subunit (ND5 homologue) of the Na+-dependent complex I from *Escherichia coli* transports Na+, *J Biol Chem.* 278, 26817-22.

82. Gemperli, A. C., Schaffitzel, C., Jakob, C. & Steuber, J. (2007) Transport of Na(+) and K (+) by an antiporter-related subunit from the Escherichia coli NADH dehydrogenase I produced in *Saccharomyces cerevisiae*, *Arch Microbiol.* 188, 509-21.

83. Gemperli, A. C., Dimroth, P. & Steuber, J. (2002) The respiratory complex I (NDH I) from *Klebsiella pneumoniae*, a sodium pump, *J Biol Chem.* 277, 33811-7.

84. Gemperli, A. C., Dimroth, P. & Steuber, J. (2003) Sodium ion cycling mediates energy coupling between complex I and ATP synthase, *Proc Natl Acad Sci U S A.* 100, 839-44.

85. Lin, P. C., Puhar, A. & Steuber, J. (2008) NADH oxidation drives respiratory Na+ transport in mitochondria from *Yarrowia lipolytica*, *Arch Microbiol.* 190, 471-80.

86. Bertsova, Y. V. & Bogachev, A. V. (2004) The origin of the sodium-dependent NADH oxidation by the respiratory chain of *Klebsiella pneumoniae*, *FEBS Lett.* 563, 207-12.

87. Stolpe, S. & Friedrich, T. (2004) The *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I) is a primary proton pump but may be capable of secondary sodium antiport, *J Biol Chem.* 279, 18377-83.

88. Drose, S., Galkin, A. & Brandt, U. (2005) Proton pumping by complex I (NADH:ubiquinone oxidoreductase) from *Yarrowia lipolytica* reconstituted into proteoliposomes, *Biochim Biophys Acta.* 1710, 87-95.

89. Batista, A. P., Marreiros, B. C., Louro, R. O. & Pereira, M. M. (2012) Study of ion translocation by respiratory complex I. A new insight using (23)Na NMR spectroscopy, *Biochim Biophys Acta.* 1817, 1810-6.

90. Batista, A. P. & Pereira, M. M. (2011) Sodium influence on energy transduction by complexes I from *Escherichia coli* and *Paracoccus denitrificans*, *Biochim Biophys Acta.* 1807, 286-92.

91. Batista, A. P., Fernandes, A. S., Louro, R. O., Steuber, J. & Pereira, M. M. (2010) Energy conservation by *Rhodothermus marinus* respiratory complex I, *Biochim Biophys Acta.* 1797, 509-15.

92. Marreiros, B. C., Batista, A. P. & Pereira, M. M. (2014) Respiratory complex I from *Escherichia coli* does not transport Na(+) in the absence of its NuoL subunit, *FEBS Lett.* 588, 4520-5.

93. Steimle, S., Bajzath, C., Dorner, K., Schulte, M., Bothe, V. & Friedrich, T. (2011) Role of subunit NuoL for proton translocation by respiratory complex I, *Biochemistry.* 50, 3386-93.

94. Roberts, P. G. & Hirst, J. (2012) The deactive form of respiratory complex I from mammalian mitochondria is a Na+/H+ antiporter, *J Biol Chem.* 287, 34743-51.

95. Maklashina, E., Kotlyar, A. B. & Cecchini, G. (2003) Active/de-active transition of respiratory complex I in bacteria, fungi, and animals, *Biochim Biophys Acta.* 1606, 95-103.

96. Soballe, B. & Poole, R. K. (1999) Microbial ubiquinones: multiple roles in respiration, gene regulation and oxidative stress management, *Microbiology.* 145 (Pt 8), 1817-30.

97. Tran, Q. H. & Unden, G. (1998) Changes in the proton potential and the cellular energetics of *Escherichia coli* during growth by aerobic and anaerobic respiration or by fermentation, *Eur J Biochem.* 251, 538-43.

98. Nicholls, D. G. & Ferguson, S. J. (2013) 3 - Quantitative Bioenergetics: The Measurement of Driving Forces in *Bioenergetics (Fourth Edition)* (Ferguson, D. G. N. J., ed) pp. 27-51, Academic Press, Boston.

99. de Graef, M. R., Alexeeva, S., Snoep, J. L. & Teixeira de Mattos, M. J. (1999) The steady-state internal redox state (NADH/NAD) reflects the external redox state and is correlated with catabolic adaptation in *Escherichia coli*, *J Bacteriol.* 181, 2351-7.

100. Schoepp-Cothenet, B., Lieutaud, C., Baymann, F., Vermeglio, A., Friedrich, T., Kramer, D. M. & Nitschke, W. (2009) Menaquinone as pool quinone in a purple bacterium, *Proc Natl Acad Sci U S A.* 106, 8549-54.

101. Pan, J. W. & Macnab, R. M. (1990) Steady-state measurements of *Escherichia coli* sodium and proton potentials at alkaline pH support the hypothesis of electrogenic antiport, *J Biol Chem.* 265, 9247-50.

I.3 Type II NADH:quinone oxidoreductase

This section is based on the following publication:

Marreiros BC, Sena FV, Sousa FM, Batista AP, Pereira MM (2016) "Type II NADH:quinone oxidoreductase family: Phylogenetic distribution, Structural diversity and Evolutionary divergences" Environmental Microbiology. DOI: 10.1111/1462-2920.13352

I.3 Type II NADH:quinone oxidoreductase

Type II NADH:quinone oxidoreductases, (NDH-2s, EC 1.6.99.3), play a central role in the energy transduction processes of the respiratory chains from fungi, protists and many prokaryotes. They are monotopic proteins with approximately 45 kDa. NDH-2s functions as an entry point for electrons into those respiratory chains and catalyzes the two electron oxidation of NADH and the reduction of quinone. NDHs-2 are also called alternative NADH:quinone oxidoreductases because they perform the same catalytic activity as respiratory Complex I, but do not contribute to the establishment of the membrane potential.

NDH-2s are present in the three domains of life, but considered to be absent in mammals. This absence in mammals and the fact of being the only respiratory NADH:quinone oxidoreductases expressed in some pathogenic microorganisms, as in the case of *P. falciparum* and *S. aureus*, or preferentially expressed in aerobiosis as for example in *E. coli* [1] made NDH-2 an attractive target for rational design of specific drugs acting at the level of the enzyme [2]. In addition, NDH-2 has also been suggested as a potential therapeutic resource for human neurodegenerative diseases, including Parkinson's disease and aging, caused by Complex I failures, because its expression restores the mitochondrial activity in animals with Complex I deficiency [3, 4].

I.3.1 NDH-2 family

NDH-2s are members of the two-Dinucleotide Binding Domains Flavoproteins (tDBDF) superfamily [5]. tDBDF members are characterized by having a NAD(P)H and a flavin binding domains, with identical structural Rossmann folds [6]. The domain closest to the N-terminal binds the flavin prosthetic group, a flavin adenine dinucleotide (FAD), and the second domain interacts with nicotinamide adenine dinucleotide (Phosphate) (NAD(P)H). All the members of the family contain FAD as prosthetic group and interact with NAD(P)H/NAD(P)$^+$, with the exception of sulfide:quinone oxidoreductase (SQR) and sulfide:flavocytochrome *c* oxidoreductase (also called flavocytochrome *c* sulfide dehydrogenase, FCSD). SQRs and FCSDs also have the two dinucleotide binding domains but the NADH binding site is structurally inaccessible owing to the presence of a loop [7-10]. Some members of the tDBDF superfamily have additional redox centers, most commonly a disulfide bridge, as for example glutathione reductases or dihydrolipoamide dehydrogenases. In fact the superfamily has also been named the flavin-disulfide reductases family. Several members of the superfamily have C-terminal extensions, which in the case of NDH-2s [11-14] (Figure I.3.1) and SQRs [7-9] constitutes the membrane interacting domain composed of two amphipathic helices.

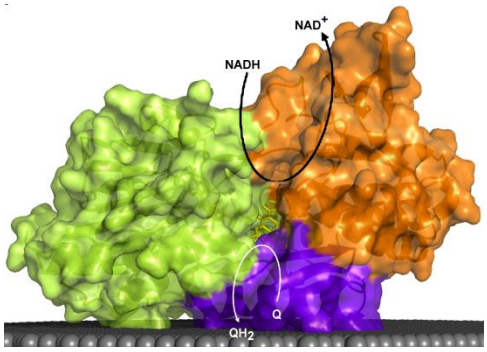


Figure I.3.1: Schematic representation of the crystal structural of NDH-2 and substrates. NDH-2 contains three domains: first dinucleotide binding domain, or FAD binding (green); second dinucleotide binding domain or NADH binding (orange); and membrane interacting domain, amphipathic helices at the C-terminal (purple). Reduction of FAD by NADH and subsequent reduction of quinone by FAD.The schematic representation is based on NDH-2s from *S. aureus* (PDB: 4XDB [14]) and *S. cerevisiae* (PDB: 4G73 [13]).

Several structural features are conserved among NDH-2 family, in *C. thermarum* NDH-2 the amino acid sequence motif AQxAxQ was proposed to characterize the quinone binding site [11], which is also present in the NDH-2s from *S. aureus* [14] and *C. cerevisiae* [12, 13]. In NDH-2s from *N. crassa* and *A. thaliana* an additional sequence insertion corresponding to an EF-hand structural motif is present (see section III.1). This motif is able to bind calcium, suggesting in this way different possibilities of protein regulation [15, 16].

I.3.2 NDH-2 crystallographic structures and constrains

The study of NDH-2s gained a new enthusiasm due to the publication of four crystallographic structures of NDH-2 from *S. cerevisiae* (PDB:4G73 [13] and PDB:4G9K [12]), *C. thermarum* (PDB: 4NWZ [11]) and *S. aureus* (PDB: 4XDB [14]). The crystal structures of yeast NDH-2 (Ndi1), obtained by two different groups, stimulated discussions on the number and location of the substrate binding sites (Figure I.3.1). The work by Iwata *et al* [12] suggested overlapping binding sites for NADH and quinone, while Feng *et al* [13] proposed separated binding sites at opposite sides of the isoalloxazine ring of FAD. The structures from the bacterial enzymes, those of *C. thermarum* [11] and *S. aureus* [14] also suggested the existence of non-overlapping binding sites for NADH and quinone, in agreement with the proposal of Feng *et al* for Ndi1. However, the structures of the bacterial enzymes were obtained in the absence of the substrates. NDH-2 from *S. aureus* was further investigated by fast kinetics and spectroscopic methods and was shown to have different binding sites for the two substrates and establishes a charge-transfer complex between NAD^+ and the flavin, which is dissociated by the quinone. This indicates the presence of a ternary complex involved in the catalytic mechanism [14] (see section III.3). Recently it was demonstrate that the NDH-2 from the bacterium *E. coli* also has two distinct substrate binding sites and it was identified a bound semi-protonated quinol as a catalytic intermediate [17]. On the other hand, one of the three NDH-2s from *M. tuberculosis* was studied by Yano and coworkers [18, 19],

whom suggested that the reaction mechanism of NDH-2 occurs through a non-classic two-site ping-pong mechanism [19].

I.3.3 NDH-2 role in cellular metabolism

Regeneration of pyridine nucleotides is central to the metabolism of most microorganisms. This is because almost all catabolic pathways involve direct or indirect transfer of electrons to NAD^+ or $NADP^+$. Their respective reduced forms NADH or NADPH may deliver electrons to respiratory chains through different NADH:quinone oxidoreductases. The simultaneous presence of genes coding for these different types of enzymes is observed in some organisms and the expression of the different enzymes depends on the organism's energetic needs and balancing of the $NADH/NAD^+$ ratio [20].

Several organisms express multiple NDH-2s; for example, in the case of fungi these enzymes are present on both sides of the mitochondrial membrane, allowing the oxidation of both the cytosolic and mitochondrial NAD(P)H (Figure I.3.2). For example *N. crassa* [21-23] NDI1 was suggested to be essential for the initial steps of germination of both sexual and asexual spores of this Ascomycota fungus [22]. *S. cerevisiae* contains two external (NDE1 and NDE2), one internal NDH-2 (Ndi1) and another one with an unknown localization (Aif1p) [24-28]. Ndi1 was observed to be essential for the yeast survival [13] and was also reported to induce apoptosis [29].

NDH-2s from protists *P. falciparum*, *T. gondii*, *T. brucei* and *C. reinhardtii* have been also investigated [30-32]. These organisms have developed different strategies to succeed in hypoxic and anoxic environments and depending on the energy metabolism demands, some relay exclusively on NDH-2 for their NADH:quinone oxidoreductase activity as in the case of *P. falciparum [30]*.

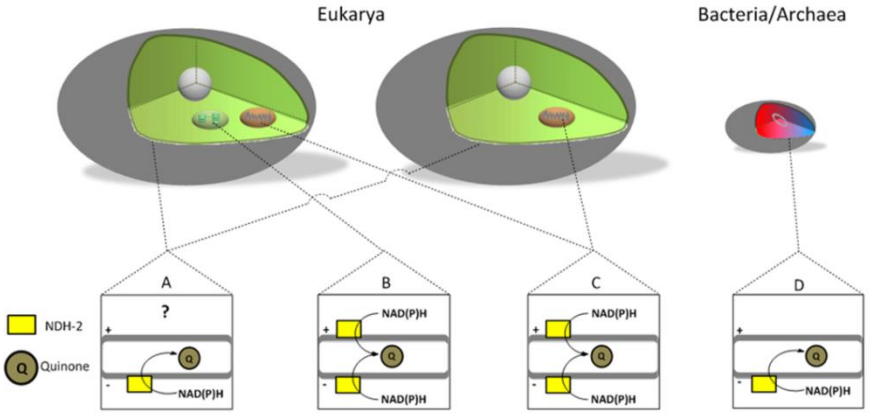


Figure I.3.2: Cellular localization of NDH-2s. Schematic representation of the localization of NDH-2s, which perform NAD(P)H:quinone oxidoreductase activity, in eukaryotic (green), bacterial (red) and archaeal (blue) cells. In Eukarya, NDH-2s may be present in the plasmatic (A), chloroplastidial (B) or mitochondrial (C) membranes and may face different sides of the membranes. In Bacteria and Archaea, NDH-2s are present in the plasmatic membrane.

I.3.4 References

1. Unden, G., Achebach, S., Holighaus, G., Tran, H. G., Wackwitz, B. & Zeuner, Y. (2002) Control of FNR function of Escherichia coli by O2 and reducing conditions, *Journal of molecular microbiology and biotechnology.* 4, 263-8.

2. Cook, G. M., Greening, C., Hards, K. & Berney, M. (2014) Energetics of pathogenic bacteria and opportunities for drug development, *Advances in microbial physiology.* 65, 1-62.

3. Schapira, A. H. (1998) Mitochondrial dysfunction in neurodegenerative disorders, *Biochimica et biophysica acta.* 1366, 225-33.

4. Dawson, T. M. & Dawson, V. L. (2003) Molecular pathways of neurodegeneration in Parkinson's disease, *Science.* 302, 819-22.

5. Ojha, S., Meng, E. C. & Babbitt, P. C. (2007) Evolution of function in the "two dinucleotide binding domains" flavoproteins, *PLoS computational biology.* 3, e121.

6. Rao, S. T. & Rossmann, M. G. (1973) Comparison of super-secondary structures in proteins, *Journal of molecular biology.* 76, 241-56.

7. Cherney, M. M., Zhang, Y., Solomonson, M., Weiner, J. H. & James, M. N. (2010) Crystal structure of sulfide:quinone oxidoreductase from *Acidithiobacillus ferrooxidans*: insights into sulfidotrophic respiration and detoxification, *Journal of molecular biology.* 398, 292-305.

8. Marcia, M., Ermler, U., Peng, G. & Michel, H. (2009) The structure of *Aquifex aeolicus* sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration, *Proceedings of the National Academy of Sciences of the United States of America.* 106, 9625-30.

9. Brito, J. A., Sousa, F. L., Stelter, M., Bandeiras, T. M., Vonrhein, C., Teixeira, M., Pereira, M. M. & Archer, M. (2009) Structural and functional insights into sulfide:quinone oxidoreductase, *Biochemistry.* 48, 5613-22.

10. Chen, Z. W., Koh, M., Van Driessche, G., Van Beeumen, J. J., Bartsch, R. G., Meyer, T. E., Cusanovich, M. A. & Mathews, F. S. (1994) The structure of flavocytochrome c sulfide dehydrogenase from a purple phototrophic bacterium, *Science.* 266, 430-2.

11. Heikal, A., Nakatani, Y., Dunn, E., Weimar, M. R., Day, C. L., Baker, E. N., Lott, J. S., Sazanov, L. A. & Cook, G. M. (2014) Structure of the bacterial type II NADH dehydrogenase: a monotopic membrane protein with an essential role in energy generation, *Molecular microbiology.* 91, 950-64.

12. Iwata, M., Lee, Y., Yamashita, T., Yagi, T., Iwata, S., Cameron, A. D. & Maher, M. J. (2012) The structure of the yeast NADH dehydrogenase (Ndi1) reveals overlapping binding sites for water- and lipid-soluble substrates, *Proceedings of the National Academy of Sciences of the United States of America.* 109, 15247-52.

13. Feng, Y., Li, W., Li, J., Wang, J., Ge, J., Xu, D., Liu, Y., Wu, K., Zeng, Q., Wu, J. W., Tian, C., Zhou, B. & Yang, M. (2012) Structural insight into the type-II mitochondrial NADH dehydrogenases, *Nature.* 491, 478-82.

14. Sena, F. V., Batista, A. P., Catarino, T., Brito, J. A., Archer, M., Viertler, M., Madl, T., Cabrita, E. J. & Pereira, M. M. (2015) Type-II NADH:quinone oxidoreductase from *Staphylococcus aureus* has two distinct binding sites and is rate limited by quinone reduction, *Molecular microbiology.* 98, 272-288.

15. Melo, A. M., Duarte, M., Moller, I. M., Prokisch, H., Dolan, P. L., Pinto, L., Nelson, M. A. & Videira, A. (2001) The external calcium-dependent NADPH dehydrogenase from *Neurospora crassa* mitochondria, *The Journal of biological chemistry.* 276, 3947-51.

16. Michalecka, A. M., Svensson, A. S., Johansson, F. I., Agius, S. C., Johanson, U., Brennicke, A., Binder, S. & Rasmusson, A. G. (2003) Arabidopsis genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light, *Plant physiology.* 133, 642-52.

17. Yano, T., Li, L. S., Weinstein, E., Teh, J. S. & Rubin, H. (2006) Steady-state kinetics and inhibitory action of antitubercular phenothiazines on *Mycobacterium tuberculosis* type-II NADH-menaquinone oxidoreductase (NDH-2), *Journal of molecular biology.* 281, 11456-11463.

18. Yano, T., Rahimian, M., Aneja, K. K., Schechter, N. M., Rubin, H. & Scott, C. P. (2014) *Mycobacterium tuberculosis* Type II NADH-Menaquinone Oxidoreductase Catalyzes Electron Transfer through a Two-Site Ping-Pong Mechanism and Has Two Quinone-Binding Sites, *Biochemistry.* 53, 1179-1190.

19. Kerscher, S., Drose, S., Zickermann, V. & Brandt, U. (2008) The three families of respiratory NADH dehydrogenases in *Bioenergetics* pp. 185-222.

20. Carneiro, P., Duarte, M. & Videira, A. (2007) The external alternative NAD(P)H dehydrogenase NDE3 is localized both in the mitochondria and in the cytoplasm of *Neurospora crassa*, *Journal of molecular biology.* 368, 1114-21.

21. Duarte, M., Peters, M., Schulte, U. & Videira, A. (2003) The internal alternative NADH dehydrogenase of *Neurospora crassa* mitochondria, *The Biochemical journal.* 371, 1005-11.

22. Weiss, H., von Jagow, G., Klingenberg, M. & Bucher, T. (1970) Characterization of *Neurospora crassa* mitochondria prepared with a grind-mill, *European journal of biochemistry / FEBS.* 14, 75-82.

23. Luttik, M. A., Overkamp, K. M., Kotter, P., de Vries, S., van Dijken, J. P. & Pronk, J. T. (1998) The *Saccharomyces cerevisiae* NDE1 and NDE2 genes encode separate mitochondrial NADH dehydrogenases catalyzing the oxidation of cytosolic NADH, *The Journal of biological chemistry.* 273, 24529-34.

24. Overkamp, K. M., Bakker, B. M., Kotter, P., van Tuijl, A., de Vries, S., van Dijken, J. P. & Pronk, J. T. (2000) In vivo analysis of the mechanisms for oxidation of cytosolic NADH by *Saccharomyces cerevisiae* mitochondria, *Journal of bacteriology.* 182, 2823-30.

25. Velazquez, I. & Pardo, J. P. (2001) Kinetic characterization of the rotenone-insensitive internal NADH: ubiquinone oxidoreductase of mitochondria from *Saccharomyces cerevisiae*, *Archives of biochemistry and biophysics.* 389, 7-14.

26. Yamashita, T., Nakamaru-Ogiso, E., Miyoshi, H., Matsuno-Yagi, A. & Yagi, T. (2007) Roles of bound quinone in the single subunit NADH-quinone oxidoreductase (Ndi1) from *Saccharomyces cerevisiae*, *The Journal of biological chemistry.* 282, 6012-20.

27. Wissing, S., Ludovico, P., Herker, E., Buttner, S., Engelhardt, S. M., Decker, T., Link, A., Proksch, A., Rodrigues, F., Corte-Real, M., Frohlich, K. U., Manns, J., Cande, C., Sigrist, S. J., Kroemer, G. & Madeo, F. (2004) An AIF orthologue regulates apoptosis in yeast, *The Journal of cell biology.* 166, 969-74.

28. Cui, Y., Zhao, S., Wu, Z., Dai, P. & Zhou, B. (2012) Mitochondrial release of the NADH dehydrogenase Ndi1 induces apoptosis in yeast, *Molecular biology of the cell.* 23, 4373-82.

29. Biagini, G. A., Viriyavejakul, P., O'Neill P, M., Bray, P. G. & Ward, S. A. (2006) Functional characterization and target validation of alternative complex I of *Plasmodium falciparum* mitochondria, *Antimicrobial agents and chemotherapy.* 50, 1841-51.

30. Fang, J. & Beattie, D. S. (2002) Novel FMN-containing rotenone-insensitive NADH dehydrogenase from *Trypanosoma brucei* mitochondria: isolation and characterization, *Biochemistry.* 41, 3065-72.

31. Lin, S. S., Kerscher, S., Saleh, A., Brandt, U., Gross, U. & Bohne, W. (2008) The *Toxoplasma gondii* type-II NADH dehydrogenase TgNDH2-I is inhibited by 1-hydroxy-2-alkyl-4(1H)quinolones, *Biochimica et biophysica acta.* 1777, 1455-62.

Part II

Results

Chapter II

Complex I family

Chapter II – Complex I family

II.1 - Complex I family: The universal adaptor

This section is based on the following publication:

Marreiros BC, Batista AP, Duarte AMS, Pereira MM (2013) “A missing link between complex I and group 4 membrane-bound [NiFe]-hydrogenases” Biochimica et Biophysica Acta (BBA)-Bioenergetics 1827 (2): 198-209. DOI:10.1016/j.bbabio.2012.09.012.

Authors' contribution:

B.C.M. participated in the design of the study, performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

II.1 - Complex I family: The universal adaptor

II.1.1 Summary

In this section we present a thorough taxonomic profile of prokaryotic Complex I family, including group 4 membrane-bound [NiFe] hydrogenases, complexes I and related enzymes (Figure 1.2.1 and 1.2.2). In addition we investigated the different gene clustering organization of such complexes. Our data show the presence of complexes related to hydrogenases but which do not contain the binding site of the catalytic [NiFe] center. These complexes, named before Ehr (Energy-converting hydrogenases related complexes) are a missing link between complex I and group 4 membrane-bound [NiFe]-hydrogenases. Based on our observations we put forward a different perspective for the relation between complex I and related complexes. In addition we discuss the evolutionary, functional and mechanistic implications of this new perspective.

II.1.2 Materials and Methods

II.1.2.1 Search of KEGG's database

Genes coding for proteins homologous to the different subunits of complex I (see Tables II.1.1 and II.1.2) were searched among the prokaryotic domains using protein BLAST (BLASTp) analysis tool running at KEGG´s (Kyoto Encyclopedia of Genes and Genomes)

database platform [1-3]. This database was selected because it contains data only on fully sequenced organisms, which is a requirement when searching for the presence or absence of genes in specific organisms.

Amino acid sequences and respective information were retrieved using a homemade script [4]. In Table II.1.1 are shown the sequences used as queries for BLASTp.

Table II.1.1: Sequences used as queries for Blastp.

Species	Complex / Subunit	KEGG ID	*Locus tag*
R. marinus	Complex I / Nqo1	rmr:Rmar_2160	Rmar_2160
R. marinus	Complex I / Nqo2	rmr:Rmar_2156	Rmar_2156
R. marinus	Complex I / Nqo3	rmr:Rmar_2162	Rmar_2162
R. marinus	Complex I / Nqo4	rmr:Rmar_2155	Rmar_2155
R. marinus	Complex I / Nqo5	rmr:Rmar_2154	Rmar_2154
R. marinus	Complex I / Nqo6	rmr:Rmar_2153	Rmar_2153
R. marinus	Complex I / Nqo7	rmr:Rmar_2152	Rmar_2152
R. marinus	Complex I / Nqo8	rmr:Rmar_2050	Rmar_2050
R. marinus	Complex I / Nqo9	rmr:Rmar_2631	Rmar_2631
R. marinus	Complex I / Nqo10	rmr:Rmar_1245	Rmar_1245
R. marinus	Complex I / Nqo11	rmr:Rmar_1246	Rmar_1246
R. marinus	Complex I / Nqo12	rmr:Rmar_1247	Rmar_1247
R. marinus	Complex I / Nqo13	rmr:Rmar_1248	Rmar_1248
R. marinus	Complex I / Nqo14	rmr:Rmar_1251	Rmar_1251
P. denitrificans	Complex I / NuoA	pde:Pden_2250	Pden_2250
P. denitrificans	Complex I / NuoJ	pde:Pden_2235	Pden_2235
P. denitrificans	Complex I / NuoK	pde:Pden_2234	Pden_2234
M. mazei	Complex I-Like / FpoF	mma:MM_0627	MM_0627
B. subtilis	Mrp / MrpA	bsu:BSU31600	BSU31600
B. subtilis	Mrp / MrpB	bsu:BSU31610	BSU31610
B. subtilis	Mrp / MrpC	bsu:BSU31620	BSU31620
B. subtilis	Mrp / MrpD	bsu:BSU31630	BSU31630
B. subtilis	Mrp / MrpE	bsu:BSU31640	BSU31640
B. subtilis	Mrp / MrpF	bsu:BSU31650	BSU31650
B. subtilis	Mrp / MrpG	bsu:BSU31660	BSU31660
B. subtilis	DUF2309 / antiporter	bsu:BSU01830	BSU01830

The selected *E-value* was ≤1 to guarantee a robust and representative sampling. 31 phyla, containing 580 classes and 1223

species (1117 species from Bacteria and 106 from Archaea) were investigated. Only one strain of each species was considered. The information used was that available by April 2012.

This search gave rise to different datasets containing the amino acid sequences, the species and gene identifications of each NuoA to N, FpoF, MrpA to G and DUF2309 related proteins.

II.1.2.2 Taxonomic profile of the universal adaptor

Each dataset obtained was processed on Microsoft Office Excel using VBA programming language. We were able to organize the data in a profile of presence and absence of each gene (*nuoA* to *N*, *fpoF*, *mrpA* to *G* and *duf2309*) by species. We consider that the universal adaptor is constituted by four subunits, NuoB, D, H and an antiporter-like subunit (L, M or N). For brevity, we will refer to the antiporter-like subunits from the universal adaptor as NuoL. The number of universal adaptors was quantified by the presence of groups of the four genes that code for their constituting subunits, which implies that one repeat of such a group of four counts as one universal adaptor, two repeated groups count as two universal adaptors, and so forth.

II.1.2.3 Taxonomic profile of hydrogenases, complex I and related complexes

Two groups of universal adaptors were considered: those belonging to hydrogenases and those present in non-hydrogenases. These groups can be distinguished based on the presence or absence of the binding site for the [NiFe] center (Figure II.1.1). Thus, a

homemade script on VBA allowed us to separate the amino acid sequences of NuoD dataset by the presence or absence of CxxC motifs of the [NiFe] center. Matching the NuoD dataset without CxxC

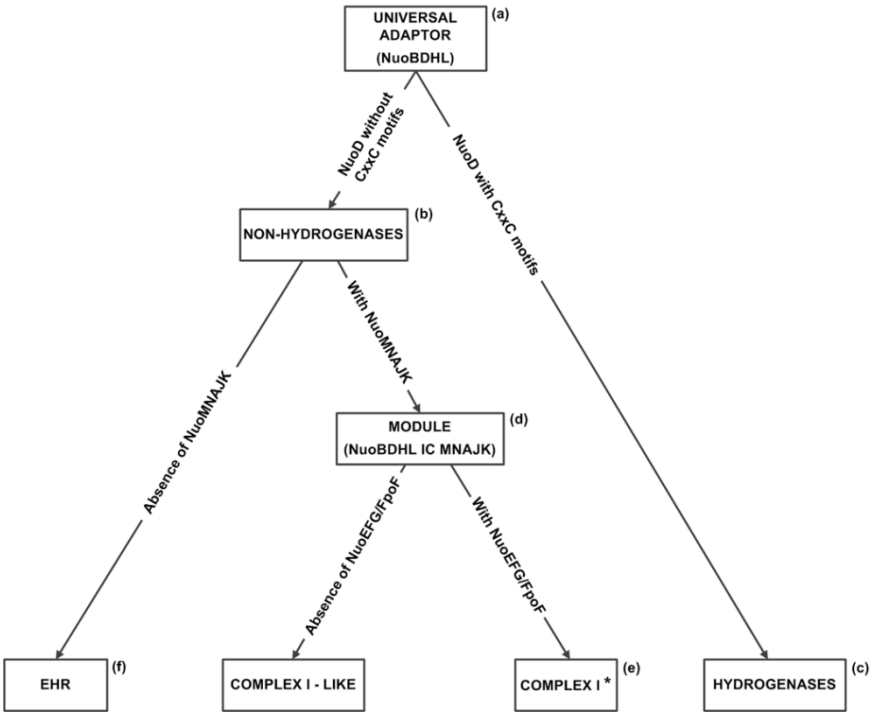


Figure II.1.1: **Graphical representation of the strategy used for the taxomonic profiles of universal adaptor:** type 4 membrane bound [NiFe] hydrogenases (hydrogenases), complex I, complex I-like enzymes and energy-converting hydrogenase related complex (Ehr). The results are presented in Table II.1.2 (letters between brackets are related with the respective columns). *Complexes I with known electron input modules (NuoEFG or FpoF).

motifs against the profile of the universal adaptor resulted in the profile of non-hydrogenases and matching the profile of NuoD with CxxC motifs against the remaining universal adaptors resulted in the

profile of hydrogenases (Figure II.1.1). Afterwards, the profile of complex I-related enzymes (module NuoBDHL IC MNAJK) and that of Ehr were obtained by matching the profiles of NuoA, C, I, J, K, M and N against non-hydrogenases profile.

Finally, matching the profiles of NuoE, F, G (in Bacteria) and FpoF (in Archaea) against Complex I-related enzymes (module NuoBDHL IC MNAJK) profile resulted in the profiles of Complex I and of Complex I-like enzymes (II.1.1).

II.1.2.4 Gene cluster organization of hydrogenases, complex I and related complexes

For each of the four final datasets (hydrogenases, complex I, complex I-like enzymes and Ehr complexes) obtained, the gene clustering organization was analyzed (see Table I.1.2). The genes coding for complex I and complex I-like enzymes correspond to 73% of the data and thus their organization was automatically analyzed using a homemade script running on VBA, which allowed grouping genes by their identification number. Two combinations were considered, namely the 14 genes cluster (*nuoA*-*N*) and the 11 genes cluster (*nuoA*-*N* without *nuoE, F* and *G*). The cases in which genes *nuoB*, *nuoC* and *nuoD* are fused were also taken into account. The other datasets (hydrogenases and Ehr complexes) corresponding to 27% of the total data were analyzed manually on KEGG´s platform, since general information on their gene clustering organization is not available.

II.1.2.5 Structural modelling

To generate the structural model of the standalone (without gene cluster) Mrp-like protein from *N. thermophilus* we used its amino acid sequence (Nther_0107). This sequence was aligned with that of NuoL from *E. coli* (via MobWeb) and the structural model was calculated with MODELLER using as template the crystallographic structure of NuoL from Complex I from *E. coli* (PDB code: 3RKO:L).

The sequence of NuoH from *E. coli* was obtained from Uniprot (C6E9S0). Sequence alignments were made using both HHPRED [5] and GenTHREADER [6]. Both methods provided the same result, which was used as input in MODELLER [7] to calculate the homology model of the NuoH from *E. coli*. As structure templates we used the crystal structures of the homologous proteins identified by the sequence alignment algorithms: NuoL (PDB code: 3RKO:L), NuoM (PDB code: 3RKO:M), NuoN (PDB code: 3RKO:N) and CysZ (PDB code: 3tx3:A). To test the robustness of the structural model, we used the protein sequence of NuoH from *E. coli* as input in the ITASSER server [8] and a second model of NuoH was generated.

Protein structure images and cartoons were generated using PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.

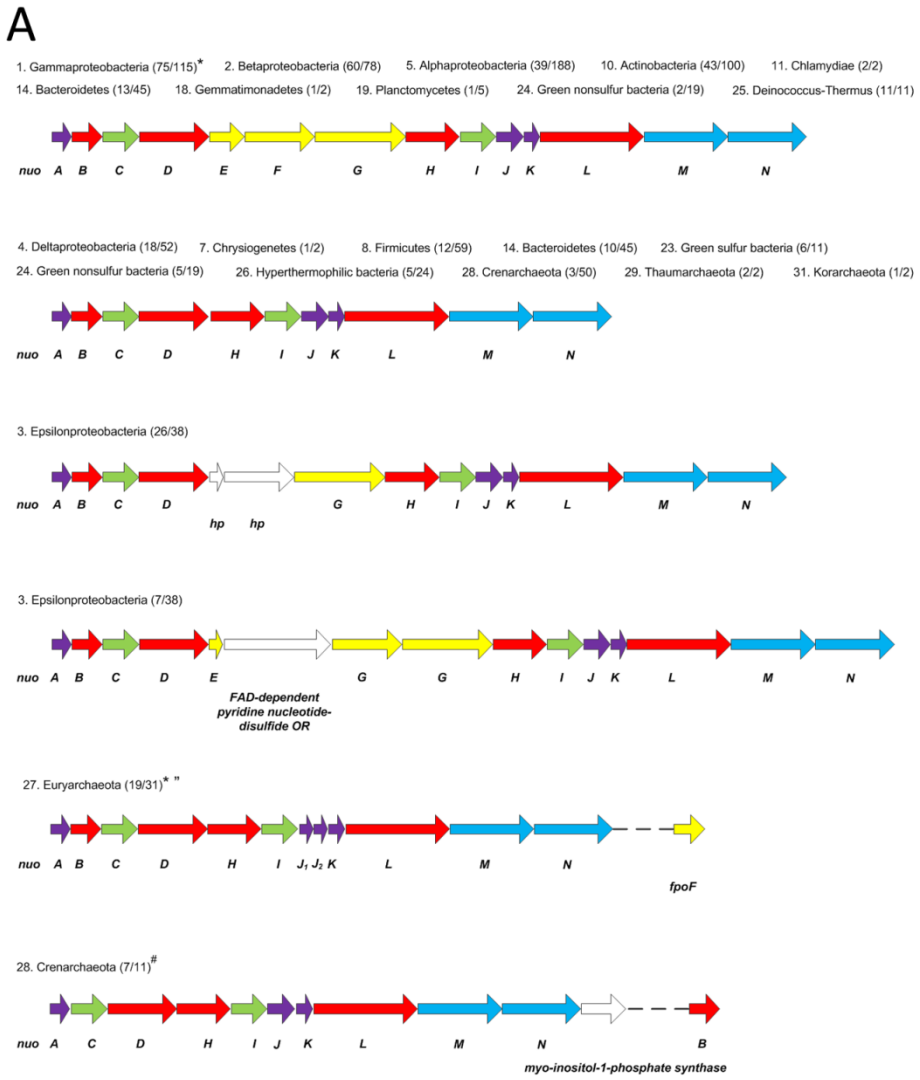
II.1.3 Results and Discussion

II.1.3.1 Taxonomic profile of hydrogenases, complex I and related complexes

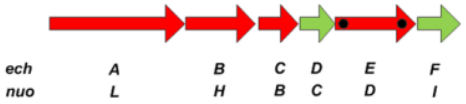
Complex I can be considered as being composed of three main parts (see section I.2). These have been named differently [9-12], but essentially they intend to define the peripheral arm in which the electron entry occurs, the membrane ion translocating machinery and a central membrane adaptor. We considered that this central membrane adaptor is composed of the soluble subunits NuoD and NuoB and the membrane subunits NuoH and NuoL, and since it is common to complex I and group 4 hydrogenases we will refer to it as the universal adaptor.

Using the sequences of NuoB, D, H and L from *R. marinus* as templates we looked for the presence of genes coding for universal adaptors in all prokaryotes with fully sequenced genome deposited in KEGG's database (Figure II.1.1). We observed that from a total of 1223 species, 831 contain at least one group of the four genes coding for the universal adaptor (Table II.1.2, column a). In total, we found 1145 groups of the four genes coding for the universal adaptor. With three exceptions these groups of genes are presented in all prokaryotic phyla (Table II.1.2).

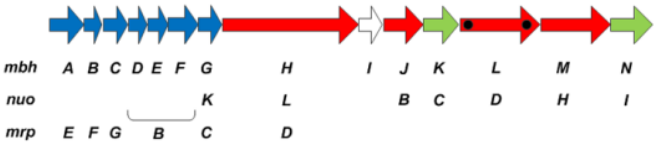
In the sequence of NuoD homologues, we searched for two CxxC motifs in which the cysteines residues provide the ligands for the [NiFe] center of hydrogenases. This allowed us to distinguish possible complexes I from possible hydrogenases (Figure II.1.1). We observed


A
1. Gammaproteobacteria (75/115)*
2. Betaproteobacteria (60/78)
5. Alphaproteobacteria (39/188)
10. Actinobacteria (43/100)
11. Chlamydiae (2/2)
14. Bacteroidetes (13/45)
18. Gemmatimonadetes (1/2)
19. Planctomycetes (1/5)
24. Green nonsulfur bacteria (2/19)
25. Deinococcus-Thermus (11/11)
nuo A B C D E F G H I J K L M N
4. Deltaproteobacteria (18/52)
7. Chrysiogenetes (1/2)
8. Firmicutes (12/59)
14. Bacteroidetes (10/45)
23. Green sulfur bacteria (6/11)
24. Green nonsulfur bacteria (5/19)
26. Hyperthermophilic bacteria (5/24)
28. Crenarchaeota (3/50)
29. Thaumarchaeota (2/2)
31. Korarchaeota (1/2)
nuo A B C D H I J K L M N
3. Epsilonproteobacteria (26/38)
nuo A B C D hp hp G H I J K L M N
3. Epsilonproteobacteria (7/38)
nuo A B C D E G G H I J K L M N
FAD-dependent pyridine nucleotide-disulfide OR
27. Euryarchaeota (19/31)* "
nuo A B C D H I J1 J2 K L M N
fpoF
28. Crenarchaeota (7/11)#
nuo A C D H I J K L M N B
myo-inositol-1-phosphate synthase
B
1. Gammaproteobacteria (25/38)
hyc B C D E F G
nuo G L H C D I B
1. Gammaproteobacteria (11/38)
3. Epsilonproteobacteria (5/9)
hyf A B C D E F G H I
nuo G L H L´ K´ M C D I B

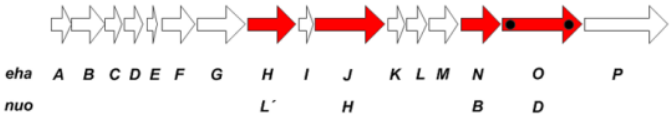
3. Epsilonproteobacteria (2/9) 4. Deltaproteobacteria (7/14)" 5. Alphaproteobacteria (4/9)* 8. Firmicutes (31/37)
10. Actinobacteria (2/4) 26. Hyperthermophilic bacteria (2/4)* 27. Euryarchaeota (8/56)



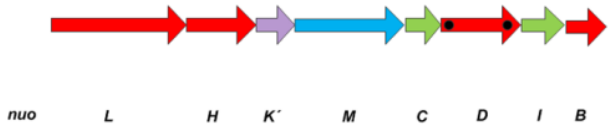
21. Synergistgetes (2/2) 27. Euryarchaeota (19/56)"



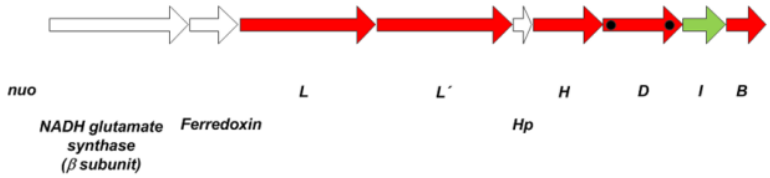
27. Euryarchaeota (19 /56)



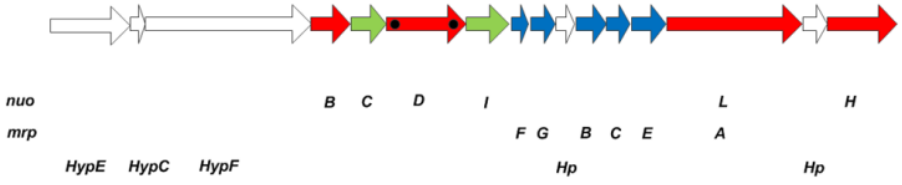
8. Firmicutes (3/37)



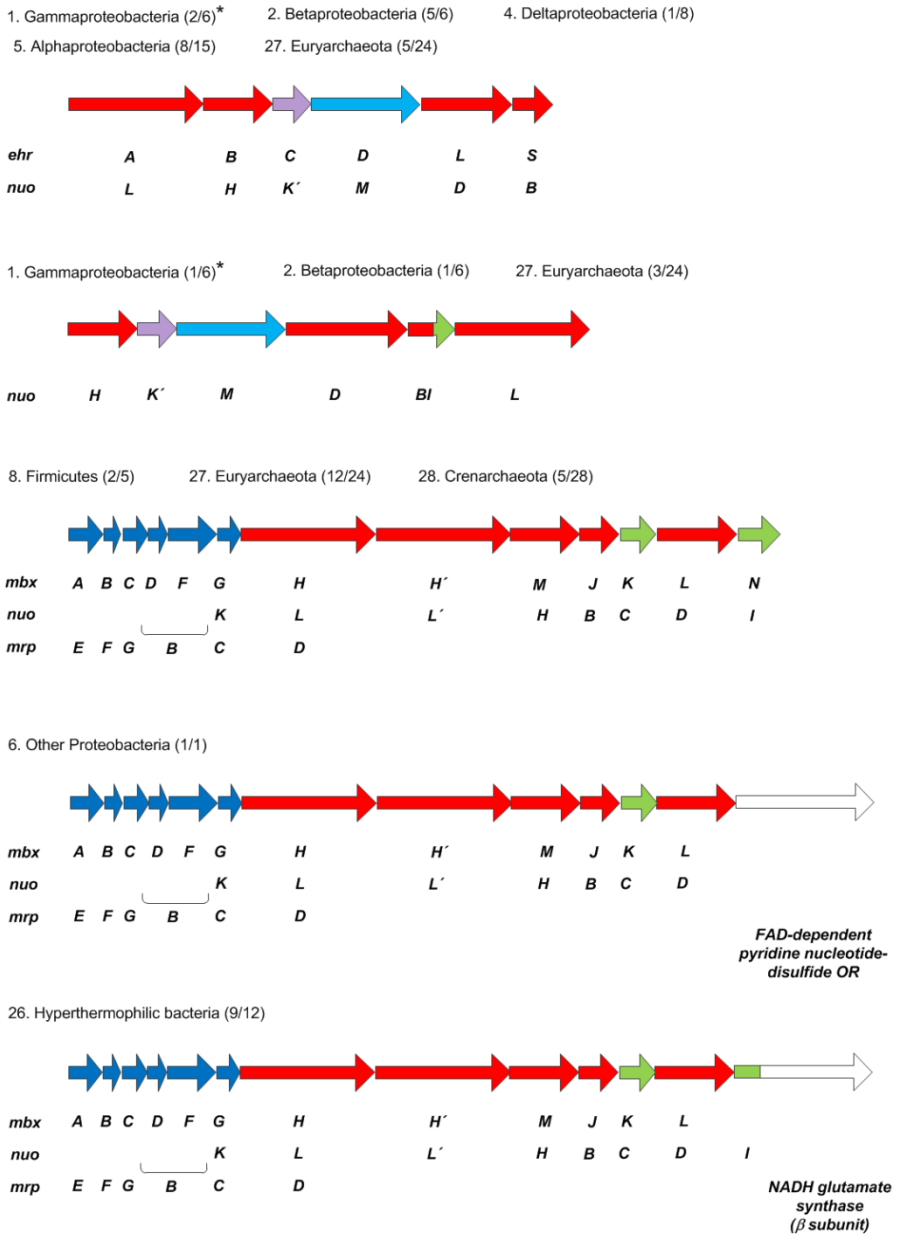
28. Crenarchaeota (4/11)



28. Crenarchaeota (5/11)



C


1. Gammaproteobacteria (2/6)*
2. Betaproteobacteria (5/6)
4. Deltaproteobacteria (1/8)
5. Alphaproteobacteria (8/15)
27. Euryarchaeota (5/24)
ehr A B C D L S
nuo L H K´ M D B
1. Gammaproteobacteria (1/6)*
2. Betaproteobacteria (1/6)
27. Euryarchaeota (3/24)
nuo H K´ M D BI L
8. Firmicutes (2/5)
27. Euryarchaeota (12/24)
28. Crenarchaeota (5/28)
mbx A B C D F G H H´ M J K L N
nuo K L L´ H B C D I
mrp E F G B C D
6. Other Proteobacteria (1/1)
mbx A B C D F G H H´ M J K L
nuo K L L´ H B C D
mrp E F G B C D
FAD-dependent pyridine nucleotide-disulfide OR
26. Hyperthermophilic bacteria (9/12)
mbx A B C D F G H H´ M J K L
nuo K L L´ H B C D I
mrp E F G B C D
NADH glutamate synthase (β subunit)

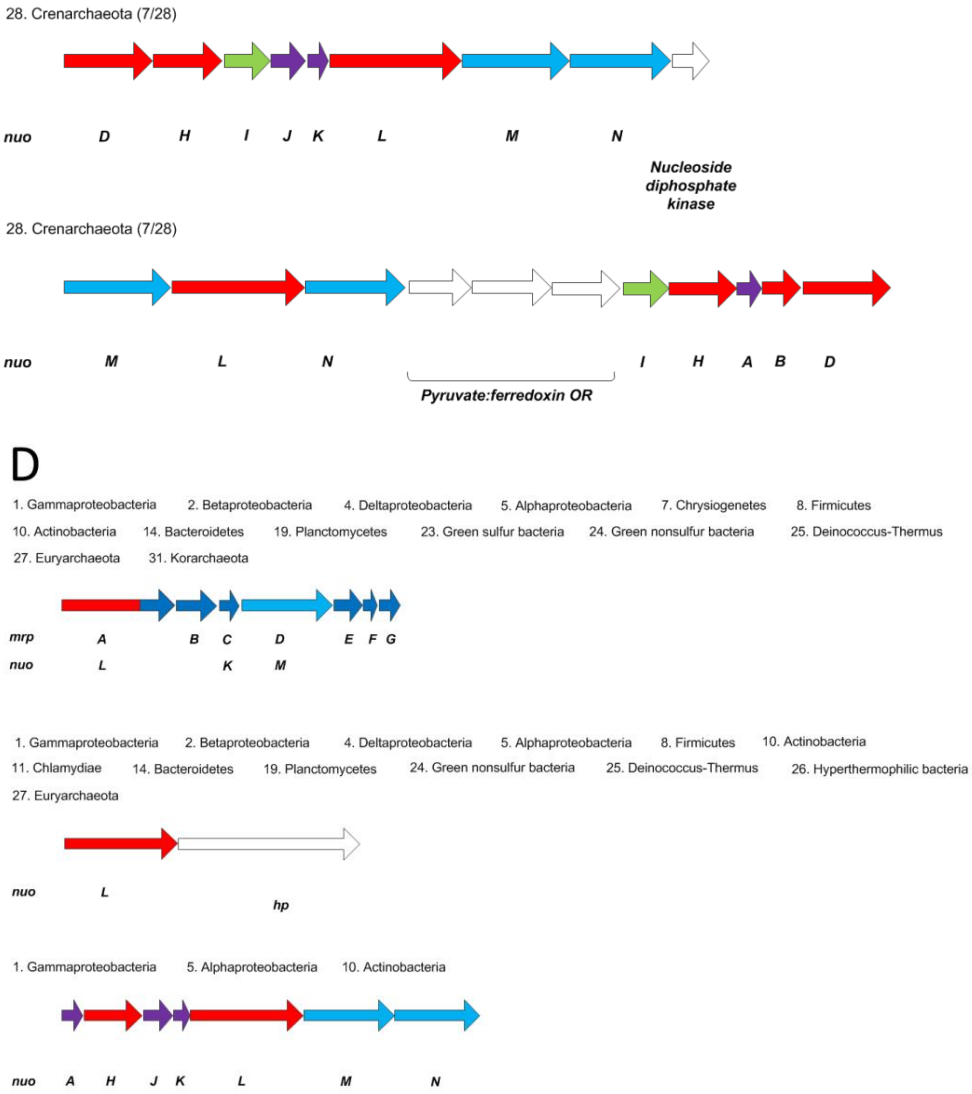


Figure II.1.2: Schematic representation of different gene cluster organizations present in the different prokaryotic phyla. A) Complex I and complex I-like enzymes. B) group 4 membrane bound [NiFe] hydrogenases. C) Energy converting hydrogenases related complexes (Ehr). D) Mrp, hypothetical protein containing Duf2309 and isolated membrane module (NuoA, H, J, K, L, M and N). The spheres represent the ligands of the NiFe center (motif: CxxC). The colour code is: red – genes coding for homologues of subunits NuoB, D, H, and L (universal adaptor); green - genes coding for homologues of subunits NuoC and NuoI; blue - genes coding for homologues of subunits NuoM and NuoN (antiporter-like subunits); purple - genes coding for homologues of subunits NuoA, J and K and yellow - genes coding for homologues of subunits NuoE, F and G (the electron input module). (Hp:

hypothetical protein; *Genes may be fused; #NuoJ may be divided in two; "Genes may be in a different order).

that 777 species contain 959 groups of genes coding for a universal adaptor without the [NiFe] center binding site, while 146 species contain 186 groups of genes with this binding site (Table II.1.2, columns b and c). The latter most probably encode membrane-bound hydrogenases.

Among the species that contain genes coding for the universal adaptor without the [NiFe] center binding site we searched for the presence of genes encoding the other subunits of complex I: NuoA, C, I, J, K, M and N. We found that 724 species have 833 groups of those genes *plus* the genes coding for the universal adaptor subunits (Table II.1.2, column d). From these species, 579 have 645 groups of genes coding for a complex I with the NuoE, F and G or FpoF input modules (Table II.1.2, column e). Most interestingly we observed the presence of 126 groups of genes, in 114 species, coding for universal adaptors without the [NiFe] center binding site, which do not have genes coding for the other subunits of complex I (Table II.1.2, column f).

Table II.1.2 clearly indicates the existence of three types of complex: group 4 hydrogenase, complex I (and related enzymes) and a membrane-bound complex related to hydrogenases, but which seems to lack the [NiFe] binding site. This type of complex is known as Ehr [13, 14].

TABLE II.1.2: Taxonomic profile of the genes coding for the subunits of the universal adaptor (a), type 4 membrane-bound [NiFe] hydrogenases (c), complex I-related enzymes (d), complex I (e), energy-converting hydrogenases related complexes (f).

				(a)			(b)			(c)			(d)			(e)			(f)		
				Universal Adaptor			Universal Adaptor (NuoD without CxxC)			Universal Adaptor (NuoD with CxxC)			Universal Adaptor (NuoD without CxxC) With IC MN AJK			Universal Adaptor (NuoD without CxxC) With IC MN AJK and EFG/FpoF			Universal Adaptor (NuoD without CxxC) Absence of IC MN AJK		
	Phylum	Class	Species	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)
Bacteria	1. Gammaproteobacteria	88	175	115	159	60	111	121	4	35	38	80	110	115	5	110	115	5	6	6	109
	2. Betaproteobacteria	44	78	77	84	1	77	84	0	0	0	77	77	78	0	77	78	0	5	6	72
	3. Epsilonproteobacteria	11	37	36	50	1	35	41	1	8	9	28	35	38	1	3	3	33	3	3	33
	4. Deltaproteobacteria	22	45	37	74	8	35	60	2	9	14	28	34	52	3	26	40	11	8	8	29
	5. Alphaproteobacteria	66	171	166	212	5	166	203	0	4	9	162	166	188	0	166	188	0	13	15	153
	6. Other proteobacteria	1	1	1	2	0	1	2	0	0	0	1	1	1	0	1	1	0	1	1	0
	7. Chrysiogenetes	1	1	1	2	0	1	2	0	0	0	1	1	2	0	1	2	0	0	0	1
	8. Firmicutes	74	249	80	101	169	59	64	21	35	37	45	55	59	25	17	20	63	5	5	75
	9. Tenericutes	6	30	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	10. Actinobacteria	67	123	81	116	42	78	112	3	3	4	78	78	100	3	77	84	4	12	12	69
	11. Chlamydiae	5	12	2	2	10	2	2	0	0	0	2	2	2	0	2	2	0	0	0	2
	12. Spirochaetes	6	27	3	3	24	3	3	0	0	0	3	3	3	0	3	3	0	0	0	3
	13. Acidobacteria	6	7	7	11	0	7	11	0	0	0	7	7	11	0	7	8	0	0	0	7
	14. Bacteroidetes	43	58	35	45	23	35	45	0	0	0	35	35	45	0	23	25	12	0	0	35
	15. Fibrobacteres	1	1	1	2	0	1	2	0	0	0	1	1	1	0	1	1	0	1	1	0
	16. Fusobacteria	5	5	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	17. Verrucomicrobia	4	4	4	8	0	4	8	0	0	0	4	4	5	0	4	5	0	2	3	2
	18. Gemmatimonadetes	1	1	1	2	0	1	2	0	0	0	1	1	2	0	1	2	0	0	0	1
	19. Planctomycetes	4	5	3	5	2	3	5	0	0	0	3	3	5	0	3	3	0	0	0	3
	20. Elusimicrobia	2	2	1	1	1	0	0	1	1	1	0	0	0	1	0	0	1	0	0	1

				(a)			(b)			(c)			(d)			(e)			(f)		
				Universal Adaptor			Universal Adaptor (NuoD without CxxC)			Universal Adaptor (NuoD with CxxC)			Universal Adaptor (NuoD without CxxC) With IC MN AJK			Universal Adaptor (NuoD without CxxC) With IC MN AJK and EFG/FpoF			Universal Adaptor (NuoD without CxxC) Absence of IC MN AJK		
	Phylum	Class	Species	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)	Species (+)	Gene (+)	Species (-)
	21. Synergistetes	3	3	2	2	1	0	0	2	2	2	0	0	0	2	0	0	2	0	0	2
	22. Cyanobacteria	12	15	15	15	0	15	15	0	0	0	15	15	15	0	0	0	15	0	0	15
	23. Green sulfur bacteria	5	10	10	11	0	10	11	0	0	0	10	10	11	0	1	1	9	0	0	10
	24. Green nonsulfur bacteria	7	12	12	22	0	12	21	0	1	1	11	12	19	0	12	18	0	2	2	10
	25. Deinococcus-Thermus	6	12	10	11	2	10	11	0	0	0	10	10	11	0	10	10	0	0	0	10
	26. Hyperthermophilic bacteria	24	33	27	40	6	27	36	0	4	4	23	18	24	9	15	17	12	12	12	15
Archaea	27. Euryarchaeota	44	70	69	111	1	50	55	19	38	56	31	31	31	38	19	19	50	23	24	46
Archaea	28. Crenarchaeota	17	31	31	50	0	30	39	1	6	11	25	11	11	20	0	0	31	21	28	10
Archaea	29. Thaumarchaeota	2	2	2	2	0	2	2	0	0	0	2	2	2	0	0	0	2	0	0	2
Archaea	30. Nanoarchaeota	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Archaea	31. Korarchaeota	2	2	2	2	0	2	2	0	0	0	2	2	2	0	0	0	2	0	0	0
TOTAL	**31**	**580**	**1223**	**831**	**1145**	**392**	**777**	**959**	**54**	**146**	**186**	**685**	**724**	**833**	**107**	**579**	**645**	**252**	**114**	**126**	**715**

II.1.3.2 Gene clusters of hydrogenases, complex I and related complexes

Table II.1.2 gives us information on the number of groups of genes coding for hydrogenases, complex I and related complexes existing in a certain genome, but does not provide any information about their possible organization into clusters. This prompted us to investigate the different organizations of gene clusters from the three types of complexes identified on Table II.1.2. Figure II.1.2 contains schematic representations of the different gene clusters we have identified present in organisms belonging to the 31 considered phyla. We could observe that from 626 groups of genes coding for complex I with NuoE, F, G as the input module, 247 are organized in gene clusters similar to that present in *E. coli* genome, either with NuoC D fused or not (Figure II.1.2, panel A). A similar gene cluster organization, but without the three genes encoding the input module proteins, is observed in 63 cases. Epsilonproteobacteria contain 38 groups of genes coding for all subunits of complex I, with the exception of genes encoding homologues of NuoE and F. We found two types of gene clusters in this phylum; one cluster containing genes coding for two hypothetical proteins, which are located between genes *nuoD* and *nuoG* and a second cluster containing an additional NuoG coding gene, a gene encoding a protein homologous to NuoE (but with a smaller size – Pfam: PF11390) and another one encoding a homologue to a protein annotated as FAD-dependent pyridine nucleotide-disulfide oxidoreductase (Pfam: PF07992). A gene coding

for FpoF, the input module of some archaeal complex I-like enzymes, is present in 19 genomes (Figure II.1.2 panel A and Table II.1.2).

We observed 56 groups of genes coding for group 4 membrane-bound hydrogenases organized as *EchABCDEF* (Figure II.1.2, panel B). Similar gene organizations to those observed in *E. coli* genome, *hyc* and *hyf* gene clusters, which code for its hydrogenases are present in other Gammaproteobacteria and in Epsilonprotobacteria (Figure II.1.2, panel B). Most interesting is the observation that the gene clusters of some hydrogenases (mbh type) contain genes coding for proteins homologous to the small subunits of Mrp Na^+/H^+ antiporters, *mrpE*, *F* and *G*, which are not present in complex I (Figure II.1.2, panels A and B).

Additionally, we observed the presence of genes coding for NuoD like subunits without the binding motifs for the [NiFe] center, which are not part of gene cluster encoding complex I or complex I-like enzymes. The genes, which code for Ehr complexes are also organized in clusters (Figure II.1.2, panel C). The simplest of these clusters contains a gene coding for an additional antiporter-like subunit and another one encoding a protein homologous to HyfE/NuoK, besides the genes coding for the universal adaptor proteins (Figure II.1.2, panel C). We also observed another gene cluster with an organization similar to that encoding for the Mbh hydrogenase, but in which the gene coding for the NuoD homologous subunit does not contain the motifs for the [NiFe] center binding. This observation has been reported before and the product of those clusters was named Mbx (membrane bound hydrogenase related). Mbx gene cluster is

structurally similar to the Mbh gene cluster but MbxL subunits do not present the binding motifs for the [NiFe] center [15] (Figure II.1.2, panel C).

In some cases we identified a gene cluster constituted by two genes: one encoding a Na^+/H^+ antiporter homologue and the second a soluble/periplasmic protein containing a domain of unknown function (DUF2309). Isolated genes coding for proteins homologues to Na^+/H^+ antiporters have also been observed and were called standalone Mrps [16]. In order to clarify the discussion we included in Figure II.1.2 the representation of the gene clusters coding for Mrp-like Na^+/H^+ antiporters (Figure II.1.2, panel D). Initial studies of Mrp antiporters considered their gene clusters to be composed of six to seven genes, but later different gene arrangements were observed, including the association of some of these genes with genes coding for group 4 hydrogenases [16, 17].

Curiously, some genomes contain gene clusters coding for the membrane arm subunits of complex I (Figure II.1.2, panel D) which seem not to have the peripheral partners. This may suggest the existence of an unknown peripheral arm, however we cannot exclude the possibility that some peripheral arm may associate with different membrane arms.

II.1.3.3 Standalone Mrp Na^+/H^+ antiporters show structural homology to the NuoL subunit

To further investigate the relation of standalone Mrp Na^+/H^+ antiporters with the antiporters-like subunits from complex I we

performed a structural homology model of the standalone Mrp from *N. thermophilus*. The sequence of the stand alone Mrp was used as query in Modweb and an alignment of this sequence with that of NuoL from *E. coli* with 25% similarity was obtained. The homology model was calculated using the structure from NuoL (PDB code: 3rko:L) as template. The final model presents 13 transmembrane α-helices organized in a similar arrangment as the NuoL subunit but do not contain a region equivalent to the 14th transmembrane helix and to helix HL. Superimposition of the standalone Mrp model with the structure of NuoL highlights the sequence and structural conservation between the two proteins, especially on amino acid residues: E144, K229 and K399 in NuoL to E144, K223 and K399 in the standalone Mrp-like protein from *N. thermophilus* (Figure II.1.3). These amino acid residues were suggested to have a key role in charge translocation by complex I [18].

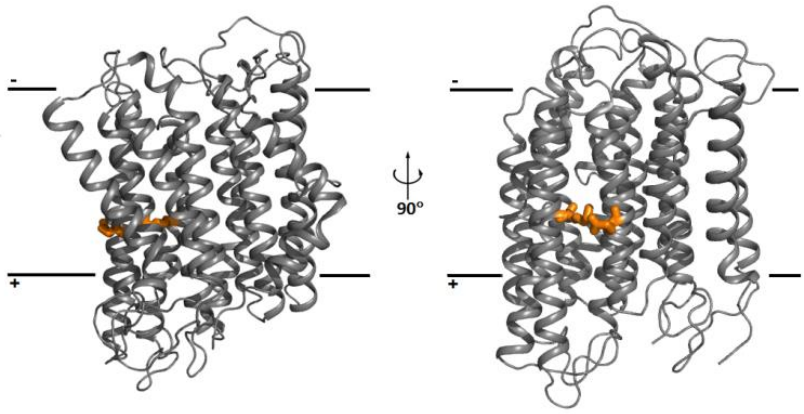


Figure II.1.3: Homology model of the Standalone Mrp from *N. thermophilus*. Side view, cytoplasmic side up. Cartoon representation of the homology model of

standalone Mrp model from *N. thermophilus*. E144 and K223 in orange stick representation.

II.1.3.4 NuoH subunit is structurally related to the antiporter-like subunits of complex I

Besides the 1179 genes coding for NuoH subunits integrated in the universal adaptor, we observed the existence of 34 other genes coding for this protein. These genes are never observed isolated and are part of a group of genes coding for the membrane subunits homologous to those of complex I, which do not have the known peripheral counterpart (Figure II.1.2, panel D).

NuoH seems to be unique to group 4 hydrogenases, Ehr complexes and complex I like proteins and is considered functionally unrelated to other known proteins [12, 19]. Curiously, blast searches for homologues only indicated other NuoH, but if we looked for paralogues within a certain genome the result was frequently one of the complex I antiporter-like subunits NuoL, NuoM or NuoN. This led us to hypothesize a relation between NuoH and these subunits. In order to investigate this possibility we have performed a structural homology model of NuoH, which was latter corroborated when the crystallographic structure of NuoH was obtained [20].

We used HHPRED and GenTHREADER to look for proteins homologous to NuoH. This strategy allowed us to identify distant sequence-structure relationships. Both algorithms showed a homology of NuoH to NuoL, NuoM and NuoN from *E. coli*, generating the same amino acid sequence alignment. This alignment was used to model the structure of NuoH from *E. coli*, using both MODELLER and ITASSER.

The resulting models were similar. The model obtained by MODELLER, which is based on the combination of several high resolution structures, shows a RMSD of 3.1 Å in the homologous regions to NuoL (Figure II.1.4). This RMSD value is within the range of values acceptable when comparing a meaningful model with its template [21]. The NuoH structural model includes 7 putative transmembrane helices. Helices 3 to 6 are structurally homologous to helices 5 to 8 of NuoL and include conserved residues, D178, T174, Y151, Y119, suggested to be part of the putative ion channel of the antiporter-like subunits [18]. Helix 3 from NuoH does not contain E144 located in homologous helix 5 from NuoL. Most interestingly helix 4 from NuoH includes E157, conserved in NuoH subunits [12] spatially close to the location where the E144 is in NuoL (blue sticks in Figure II.1.4). In helix 6 from NuoH we can also identify a helix kink structurally homologous to the kink observed in Helix 7 from NuoL and where a highly conserved aspartate residue of NuoH [12] is located (D213).

Not aiming at providing a structure for NuoH the model supports our hypothesis for a structural relation between NuoH and the antiporter-like subunits of complex I (NuoL/M/N).

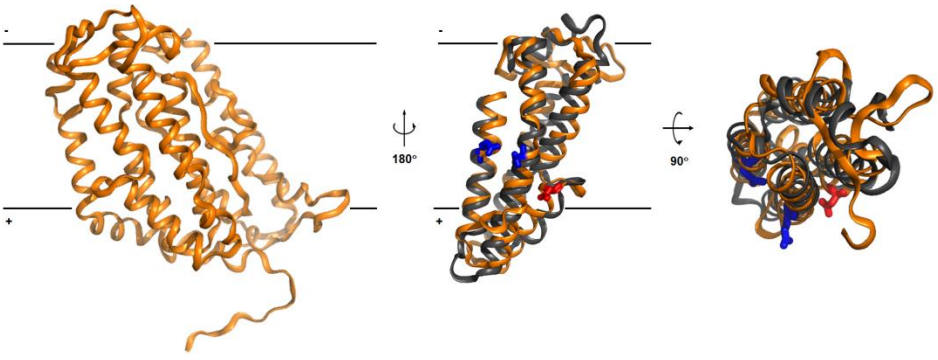


Figure II.1.4: Homology model of the NuoH from *E. coli*. Side view, cytoplasmic side up. Left: Homology model of the NuoH from E. coli in cartoon representation. Middle: superimposition of the homologous region of NuoH (in gold) to the NuoL (from 3rko:L) from *E. coli* (in dark grey), in blue sticks the E144 (from NuoL) and the E157 (from NuoH). In red sticks is represented the D213 from NuoH. Right: Top view from the homologous region of NuoH (couloring as before).

II.1.3.5 Ehr complexes, a new piece for the puzzle of hydrogenases and complex I

Based on the data presented in Table II.1.2 and on the organizations of the gene clusters (Figure II.1.2), we identified membrane complexes similar to membrane-bound [NiFe] hydrogenases, which do not have the motifs for the binding of the [NiFe] center. The possible existence of such complexes had already been reported [13-15], and were named Ehr [13, 14].

As described above, the simplest gene clusters coding for Ehr complexes, such as those from Proteobacteria and Euryarchaeota, should code for two peripheral subunits homologous to NuoD and NuoB and four membrane proteins, one homologous to NuoH, two equivalent to NuoL and one homologous to HyfE (whose C-terminus is similar to NuoK) of Hyf hydrogenase from *E. coli* (Figures 3.1.1 panel A and Figure II.1.2 panel C).

In Firmicutes, hyperthermophilic Bacteria and Archaea the Ehr complex may be composed of additional small membrane subunits homologous to MprB, E, F and G. In this case the presence of subunits homologous to NuoC and NuoI is also expected (Figure II.1.2 panel B and 3.1.2 panel C). Interestingly, the NuoL-like subunit from hyperthermophilic bacteria seems to be fused to a domain similar to the small subunit of the glutamate synthase (Pfam: PF07992;

PF12798). This is a flavoprotein that interacts with NADH [22, 23] and is thus most probably the input module of these systems (Figure II.1.2, panel C).

The so-called Mbx from the archaeal order Thermococcales, which includes *P. furiosus* and *Thermococcus (T.) kodakarensis* have been genetically characterized and were suggested to function as ferredoxin:$NADP^+$ oxidoreductases [24, 25]. However, was reported recently that NuoD subunit from Mbx complexes have an active cysteine pair instead of the two CxxC motifs that bind the [NiFe] cluster, the catalytic center of hydrogenases complexes. Therefore ferredoxin:sulfur oxidoreductase activity, coupled to ion transport was also proposed for Mbx [26]. It can be hypothesized that these complexes couple peripheral catalytic reactions to charge translocation, without involving hydrogen as substrate or product. Since the genes coding for Mbx complexes are very similar to those coding for the Mbh hydrogenase, which was shown to perform ion translocation, Mbx complexes most probably also do so [24]. Curiously, in these complexes the subunit homologous to NuoB and NuoI also contains the binding motifs for the iron-sulfur centers. The presence of such centers allows electron conduction up to the membrane surface and thus we hypothesize that the catalytic reaction of at least some of these new complexes may involve quinones.

Ehr complexes are here recognized as members of the large family of proteins including hydrogenases, complexes I and complex I – like enzymes.

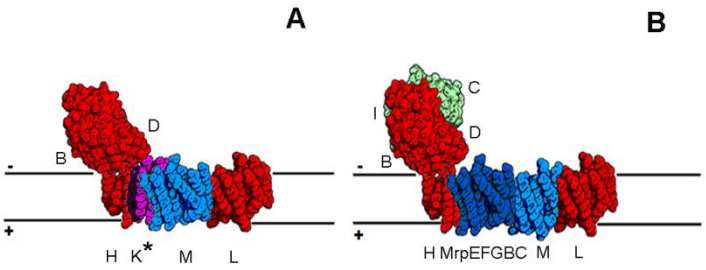


Figure II.1.4: Schematic representation of energy-converting hydrogenase related complexes (Ehr). A) energy-converting hydrogenase related. B) membrane bound hydrogenase related (Mbx). The colour code is: red - homologous subunits to NuoB, D, H, and L (Ehr: EhrS, EhrL, EhrB and EhrA; Mbx: MbxJ, MbxL, MbxM and MbxH); green - homologous subunits to NuoC and NuoI (Mbx: MbxK and MbxN); blue - homologous subunit to NuoM/N and MrpE, F, G, B (Ehr: EhrD; Mbx: MbxH´, MbxA to D and MbxF) and purple homologous subunit to NuoK (Ehr: EhrC; Mbx: MbxG) (K*: Protein is homologous to the C-terminus of NuoK) (see also Figures 1.2.1 and 1.2.2).

II.1.3.6 Evolution of group 4 membrane-bound [NiFe] hydrogenases and complex I

As reported before a close evolutionary relationship between group 4 membrane-bound [NiFe] hydrogenases and complex I has long been assumed [12, 14, 19, 27-29]. Common to the two complexes are four subunits (NuoB, D, H and L) that we nominated as the universal adaptor. Two of the subunits are peripheral (NuoB and D), facing the cytoplasm whereas the other two are integral membrane proteins (NuoH and L). Performing a taxonomic profile of the universal adaptor we observed that, with the exceptions of three phyla, the genes that code for the four proteins are present in all other prokaryotic phyla (Table II.1.2). Furthermore, besides complex I and group 4 membrane-bound [NiFe] hydrogenases, Ehr complexes also contain that same universal adaptor (Table II.1.2). The genes coding for all these complexes are mainly organized in clusters (Figure II.1.2).

II.1.3.7 Origin of the Universal Adaptor

The subunits NuoB and NuoD are homologous to the small and large subunits of soluble [NiFe] hydrogenases [12, 28]. These soluble hydrogenases are capable of performing the reduction of protons to hydrogen without the involvement of any other protein. So why did a soluble enzyme became membrane-bound and associated with proteins homologous to NuoH and Na^+/H^+ antiporter-like, as it happens in the case of Ech? In order to answer this question we may analyze which evolutionary constrains or advantages existed for a soluble enzyme to become membrane bound. We anticipate three possibilities: 1) the catalytic reaction is not thermodynamically favourable and needs energy, such as the dissipation of the transmembrane difference of electrochemical potential, 2) the catalytic reaction is thermodynamically favourable and the free energy may be transduced and conserved in the form of the transmembrane difference of electrochemical potential or 3) the catalytic reaction evolves to have a membrane soluble substrate.

Membrane-bound hydrogenases kept their catalytic activities and the enzymes from the archaea *M. mazei* (Ech) and *P. furiosus* (Mbh) were shown to perform energy conservation [30, 31]. Thus, we hypothesize that the initial advantage for the membrane association of hydrogenases was to benefit from or contribute to the membrane potential.

Since the catalytic activity occurs at the peripheral subunits the contribution to the membrane electrochemical potential cannot occur

by a charge separation mechanism, thus charge translocation across the membrane must be involved. In this way it is not surprising that hydrogenases have recruited an antiporter-like protein.

The origin of NuoH (and the reason for being recruited) is not so clear. This protein is considered functionally unrelated to any other known protein [12, 19] and seems to be unique to complex I, Ehr complexes and group 4 membrane-bound [NiFe] hydrogenases. However, our incident search for paralogues of NuoH and its structural homology model (Figure II.1.4) lead us to propose a relation between NuoH and the antiporter-like subunits (NuoL, NuoM and NuoN), which may reflect an evolutionary link between the Na^+/H^+ antiporter-like subunit and NuoH, the two membrane subunits of the universal adaptor. It can be hypothesized that NuoH arose by a gene duplication of a Na^+/H^+ antiporter-like subunit, which then became structurally and functionally specialized in order to allow efficient coupling between the cytoplasmatic reaction and the membrane charge translocation. Our results were confirmed later to be correct with the resolution of the entire crystallographic structure of complex I from *T. thermophilus* [20].

II.1.3.8 Origin of the additional subunits, Peripheral subunits

The peripheral arm of complex I and of most membrane-bound [NiFe] hydrogenases is constituted, besides the two subunits NuoD and NuoB, by two other subunits, NuoC and NuoI. The latter is similar to soluble ferredoxins, while the role of the former is less clear; the

protein does not contain any cofactor and in several cases it is fused to NuoD.

The classical complex I receives electrons from NADH through subunits NuoE, F and G. The need for all these subunits is not known and especially intriguing is the conservation of the C-terminus of NuoG. As mentioned in the Introduction, this domain is homologous to formate dehydrogenase, although the binding ligands of the catalytic center, a molybdopterin, are not present. The C-terminus of NuoG is not involved in electron transfer and is a large domain. The NuoE, F and G composition is similar to the soluble NAD^+ reducing formate dehydrogenase from *Ralstonia eutropha* [32]. The same composition, but without the domain equivalent to that of the C-terminus of NuoG is observed in the so-called bidirectional NAD^+ reducing hydrogenases (Hox) from cyanobacteria [33].

It can be speculated that NuoF could directly interact with NuoB, thus the need for the long electron conducting wire composed of the iron-sulfur centers is again intriguing. The constitution of the peripheral arm of complex I could have been the result of a random evolutionary assembling of pre-existing modules. However, the synthesis of the peripheral subunits is such an energetically expensive process which most probable exerts evolutional pressure to change complex I composition. In this way, we hypothesize that the composition of the peripheral arm of complex I is mechanistically relevant.

II.1.3.9 Origin of the additional subunits, Membrane subunits

Complex I and more complex membrane-bound [NiFe] hydrogenases evolved by one or several duplications of the gene coding for Na^+/H^+ antiporter subunit and recruitment of other small membrane proteins, such as those homologous to HyfE/NuoK (these are in turn homologous to MrpC, Figure II.1.2, panel B). The replication of Na^+/H^+ antiporter subunit would confer additional ion translocating sites and would possibly allow a higher efficiency in energy conservation. The function of the small subunits is still enigmatic. Most interesting is the observation that the small subunits in membrane-bound hydrogenases, are homologous to the Mrp small subunits, MrpE, F and G (Figure II.1.2). The association of some of the genes coding for Mrp subunits to hydrogenase or hydrogenase-like genes has been indicated before [34].

II.1.3.10 Complex I is not the evolutionary result of the association of a soluble NAD^+ reducing hydrogenase with a Mrp antiporter

Contrary to what has been proposed [19] complex I does not seem to be the evolutionary result of the association of a soluble NAD^+ reducing hydrogenase with a multisubunit Mrp antiporter complex. Complex I seems to result from evolutionary pressures on soluble hydrogenases to become membrane associated: soluble hydrogenases associated with one or two antiporter-like subunits. One of these antiporter subunits evolved as NuoH. Later the genes coding for the other antiporter duplicated possibly to increase energy transduction efficiency. Eventually this new membrane arm dissociated from the

complex giving rise to the multisubunit Mrp Na^+/H^+ antiporter complexes. In this way we suggest that multisubunit Mrp complexes are a result of the evolution of the membrane arm of hydrogenases or related complexes.

Ehr complexes, most probably arose from hydrogenases by the loss of their [NiFe] center binding site. Two types of Ehr complexes may be considered; one close to Ech and the other related to Mbh hydrogenases. In fact it has been suggested that hydrogenases and this type of complexes were acquired by different bacteria by three gene transfer events [35]. Two of these involved the transfer of the *ech* gene cluster independently from the *Methanosarcina* genus to some species from the Deltaproteobacteria and Firmicutes phyla. The other was the transfer of the *mbx* gene cluster from the order Thermococcales to species of hyperthermophilic bacteria [35]. These observations suggest that the loss of the [NiFe] binding site occurred more than once, which means that evolution of these complexes was not linear, but comprised multiple and parallel events.

The catalytic reaction of Ehr complexes is not known, however it is tempting to speculate that some may have acquired a quinone interacting site. Complex I related enzymes may have evolved from these quinone interacting Ehr complexes or directly from hydrogenases. Most interesting is the observation, in the genome of *Pelobacter propionicus* of a complex (Ppro_0587 - Ppro_0598) composed by the same subunits as those present in complex I-like enzymes, but whose NuoD homologous subunit contains the binding motif of the [NiFe] center.

II.1.3.11 Functional implications

Our analysis led us to several observations that allow a new discussion on the mechanism of not only complex I and complex I-like enzymes, but also of group 4 membrane-bound [NiFe] hydrogenases and Ehr complexes. All these complexes have in common the subunits composing the universal adaptor module. This module is most probably structural and functionally involved in the coupling of the two activities performed by the complexes, *i.e.* the coupling between the peripheral catalytic activity and the membrane charge translocation.

Initially it was proposed that Mbx complexes have ferredoxin:$NADP^+$ oxidoreductase activity [24, 25]. This reaction does not involve the presence of a catalytic site close to the membrane surface, therefore suggesting the action of long range conformational changes for the coupling of the catalytic reaction to charge translocation. Furthermore, we have mentioned that the length of the peripheral arm of complex I is intriguing, since in theory a NuoF-like protein could directly interact with NuoB. We then hypothesize that, considering its size, shape and its relative position to the membrane arm the peripheral arm total assembly may be functionally relevant to the coupling mechanism and not merely be a scaffold for the electron transfer wire (which should not be present in Ehr complexes). We suggest that the peripheral arm may work as a *mechanical lever* and as such a small effect at its top may have a large repercussion at its base.
The discussion presented here has also repercussions in terms of the ion translocation machinery including that of the multisubunit Mrp

antiporter complexes. The homology between the antiporter-like subunits (NuoL, M and N) of complex I and subunits of the Mrp Na^+/H^+ antiporter complexes, besides the amino acid sequence similarity is now also supported by structural analyses [18, 36]. Until recently Mrps were thought to be functional only within a complex [17]. However the standalone Mrp-like protein from *N. thermophilus* was shown to have Na^+/H^+ antiporter activity [37]. The homology model of this Mrp (Figure II.1.3) highlights the similarity of this protein to NuoL homologues, especially in the case of amino acid residues considered key elements for charge translocation.

II.1.4 Conclusion

The understanding of the modularity of complex I, group 4 membrane-bound [NiFe] hydrogenases and Ehr complexes will help to unravel the coupling mechanism between catalytic reactions and ion translocation by this family of proteins. We hypothesize, based on their close relationship that the three types of complexes may have a similar coupling mechanism. Membrane-bound hydrogenases, complex I and complex I like enzymes have a catalytic center ([NiFe] or quinone binding site) close to the membrane surface. By contrast some Ehr complexes may perform catalytic reactions distant from the membrane surface. In this case the coupling mechanism should be based on long range conformational changes, and not on possible conformational changes at the catalytic sites close to the membrane

surface. This could also provide hints for an explanation for the need of a long peripheral arm in complex I.

Our analysis led us to define the universal adaptor, which is common to all enzymes constituting the Complex I family. In this way, we further investigate the type of the Na^+/H^+ antiporter-like subunit (NuoL, NuoM or NuoN) that is part of the universal adaptor (see section II.2).

II.1.5 Acknowledgements

We thank Ricardo O. Louro for all the discussions, especially on molecular levers, and critical reading of the manuscript. We thank Filipa L. Sousa for the scripts used to retrieve protein sequences from KEGG's database and Cláudio M.Soares for technical discussions. We also acknowledge Miguel Teixeira, Patrícia N. Refojo and Inês C. Pereira for critical reading the manuscript and discussions. A.P.B. and A.M.D. are recipients of grants from Fundação para a Ciência e a Tecnologia (SFRH/BPD/80741/2011 and SFRH/BPD/78075/2011, respectively). The project and B.C.M. fellowship were funded by Fundação para a Ciência e a Tecnologia (PTDC/QUI-BIQ/100302/2008 to M.M.P.). The work was supported by Fundação para a Ciência e a Tecnologia through grant # PEst-OE/EQB/LA0004/2011.

II.1.6 References

1. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. & Yamanishi, Y. (2008) KEGG for linking genomes to life and the environment, *Nucleic Acids Res.* 36, D480-4.

2. Kanehisa, M. & Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.* 28, 27-30.

3. Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. & Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG, *Nucleic Acids Res.* 34, D354-7.

4. Sousa, F. L., Alves, R. J., Pereira-Leal, J. B., Teixeira, M. & Pereira, M. M. (2011) A bioinformatics classifier and database for heme-copper oxygen reductases, *PLoS One.* 6, e19117.

5. Soding, J. (2005) Protein homology detection by HMM-HMM comparison, *Bioinformatics.* 21, 951-60.

6. McGuffin, L. J. & Jones, D. T. (2003) Improvement of the GenTHREADER method for genomic fold recognition, *Bioinformatics.* 19, 874-81.

7. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. & Karplus, M. (1995) Evaluation of comparative protein modeling by MODELLER, *Proteins.* 23, 318-26.

8. Roy, A., Kucukural, A. & Zhang, Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction, *Nat Protoc.* 5, 725-38.

9. Mathiesen, C. & Hagerhall, C. (2003) The 'antiporter module' of respiratory chain complex I includes the MrpC/NuoK subunit -- a revision of the modular evolution scheme, *FEBS Lett.* 549, 7-13.

10. Friedrich, T. (2001) Complex I: a chimaera of a redox and conformation-driven proton pump?, *J Bioenerg Biomembr.* 33, 169-77.

11. Kerscher, S., Drose, S., Zickermann, V. & Brandt, U. (2008) The three families of respiratory NADH dehydrogenases, *Results Probl Cell Differ.* 45, 185-222.

12. Efremov, R. G. & Sazanov, L. A. (2012) The coupling mechanism of respiratory complex I - A structural and evolutionary perspective, *Biochim Biophys Acta*, doi:10.1016/j.bbabio.2012.02.015.

13. Coppi, M. V. (2005) The hydrogenases of *Geobacter sulfurreducens*: a comparative genomic perspective, *Microbiology.* 151, 1239-54.

14. Vignais, P. M. & Billoud, B. (2007) Occurrence, classification, and biological function of hydrogenases: an overview, *Chem Rev.* 107, 4206-72.

15. Silva, P. J., van den Ban, E. C., Wassink, H., Haaker, H., de Castro, B., Robb, F. T. & Hagen, W. R. (2000) Enzymes of hydrogen metabolism in *Pyrococcus furiosus*, *Eur J Biochem.* 267, 6541-51.

16. Krulwich, T. A., Hicks, D. B. & Ito, M. (2009) Cation/proton antiporter complements of bacteria: why so large and diverse?, *Mol Microbiol.* 74, 257-60.

17. Swartz, T. H., Ikewada, S., Ishikawa, O., Ito, M. & Krulwich, T. A. (2005) The Mrp system: a giant among monovalent cation/proton antiporters?, *Extremophiles.* 9, 345-54.

18. Efremov, R. G. & Sazanov, L. A. (2011) Structure of the membrane domain of respiratory complex I, *Nature.* 476, 414-20.

19. Moparthi, V. K. & Hagerhall, C. (2011) The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits, *J Mol Evol.* 72, 484-97.

20. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. (2013) Crystal structure of the entire respiratory complex I, *Nature.* 494, 443-8.

21. Marks, D. S., Colwell, L. J., Sheridan, R., Hopf, T. A., Pagnani, A., Zecchina, R. & Sander, C. (2011) Protein 3D structure computed from evolutionary sequence variation, *PLoS One.* 6, e28766.

22. Suzuki, A. & Knaff, D. B. (2005) Glutamate synthase: structural, mechanistic and regulatory properties, and role in the amino acid metabolism, *Photosynth Res.* 83, 191-217.

23. Vanoni, M. A. & Curti, B. (2008) Structure-function studies of glutamate synthases: a class of self-regulated iron-sulfur flavoenzymes essential for nitrogen assimilation, *IUBMB Life.* 60, 287-300.

24. Kanai, T., Matsuoka, R., Beppu, H., Nakajima, A., Okada, Y., Atomi, H. & Imanaka, T. (2011) Distinct physiological roles of the three [NiFe]-hydrogenase orthologs in the hyperthermophilic archaeon *Thermococcus kodakarensis*, *J Bacteriol.* 193, 3109-16.

25. Bridger, S. L., Clarkson, S. M., Stirrett, K., DeBarry, M. B., Lipscomb, G. L., Schut, G. J., Westpheling, J., Scott, R. A. & Adams, M. W. (2011) Deletion strains reveal metabolic roles for key elemental sulfur-responsive proteins in *Pyrococcus furiosus*, *J Bacteriol.* 193, 6498-504.

26. Schut, G. J., Zadvornyy, O., Wu, C. H., Peters, J. W., Boyd, E. S. & Adams, M. W. (2016) The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor, *Biochim Biophys Acta*.

27. Hedderich, R. & Forzi, L. (2005) Energy-converting [NiFe] hydrogenases: more than just H2 activation, *J Mol Microbiol Biotechnol.* 10, 92-104.

28. Friedrich, T. & Weiss, H. (1997) Modular evolution of the respiratory NADH:ubiquinone oxidoreductase and the origin of its modules, *J Theor Biol.* 187, 529-40.

29. Hedderich, R. (2004) Energy-converting [NiFe] hydrogenases from archaea and extremophiles: ancestors of complex I, *J Bioenerg Biomembr.* 36, 65-75.

30. Sapra, R., Bagramyan, K. & Adams, M. W. (2003) A simple energy-conserving system: proton reduction coupled to proton translocation, *Proc Natl Acad Sci U S A.* 100, 7545-50.

31. Welte, C., Kratzer, C. & Deppenmeier, U. (2010) Involvement of Ech hydrogenase in energy conservation of *Methanosarcina mazei*, *FEBS J.* 277, 3396-403.

32. Horch, M., Lauterbach, L., Lenz, O., Hildebrandt, P. & Zebger, I. (2012) NAD(H)-coupled hydrogen cycling - structure-function relationships of bidirectional [NiFe] hydrogenases, *FEBS Lett.* 586, 545-56.

33. Tamagnini, P., Leitao, E., Oliveira, P., Ferreira, D., Pinto, F., Harris, D. J., Heidorn, T. & Lindblad, P. (2007) Cyanobacterial hydrogenases: diversity, regulation and applications, *FEMS Microbiol Rev.* 31, 692-720.

34. Lim, J. K., Kang, S. G., Lebedinsky, A. V., Lee, J. H. & Lee, H. S. (2010) Identification of a novel class of membrane-bound [NiFe]-hydrogenases in *Thermococcus onnurineus* NA1 by in silico analysis, *Appl Environ Microbiol.* 76, 6286-9.

35. Calteau, A., Gouy, M. & Perriere, G. (2005) Horizontal transfer of two operons coding for hydrogenases between bacteria and archaea, *J Mol Evol.* 60, 557-65.

36. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I, *Nature.* 465, 441-5.

37. Mesbah, N. M., Cook, G. M. & Wiegel, J. (2009) The halophilic alkalithermophile *Natranaerobius thermophilus* adapts to multiple environmental extremes using a large repertoire of Na(K)/H antiporters, *Mol Microbiol.* 74, 270-81.

II.2 - Complex I family: Na^+/H^+ antiporter subunit from the Universal adaptor

This section is based on the following publication:

Batista AP, **Marreiros BC**, Pereira MM (2013) "The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes" Biological chemistry 394 (5), 659-666. DOI:10.1515/hsz-2012-0342

Authors' contribution:

B.C.M. participated in the design of the study, performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

II.2 - Complex I family: Na^+/H^+ antiporter subunit from the Universal adaptor

II.2.1 Summary

We investigated the long recognized relationship between Complexes I and group 4 [NiFe] hydrogenases and we established the so-called energy-converting hydrogenase related complexes (Ehr) as new members of the Complex I family. We also observed that four subunits, homologues to NuoB, D, H and a Na^+/H^+ antiporter-like subunit are common to the members of the family. We designated this common group of subunits as the universal adaptor. Taking into account the similarity among the Na^+/H^+ antiporter-like subunits of complex I (NuoL, NuoM and NuoN) and the unique structural characteristic of the long amphipathic α-helix part of NuoL, the nature of the antiporter-like subunit of the universal adaptor was questioned. Thus, in this work we further explored the properties of the universal adaptor, investigating which antiporter-like subunit is part of the universal adaptor. We observed that the universal adaptor contains an antiporter-like subunit with a long amphipathic α-helix as NuoL. Consequently, the long helix is also a common denominator that has been conserved in all members of the family. Such conservation surely reflects the key role of such helix in the energy transduction mechanism of this family of enzymes.

II.2.2 Materials and Methods

II.2.2.1 Sequence analyses

The proteins selected were: NuoL, NuoM, NuoN, HycC, HyfB, HyfD and HyfF from *E. coli* K-12 MG1655 (b2278, b2277, b2276, b2723, b2482, b2484, b2486), EchA, EhrA and EhrD from *Methanosarcina mazei* (MM_2320, MM_1059, MM_1062), MbhH+MbhI, MbxH and MbxH´ from *Pyrococcus furiosus* DSM 3638 (PF1430 + PF1431, PF1447, PF1446). Amino acid sequences of the antiporter-like subunits were aligned using ClustalX v.2.0. [1]. Default parameters of ClustalX were used, as no significant differences were observed with different parameter combinations. Protein weight matrix Gonnet, with Gap Opening 10 and Gap Extension 0.2 was used for multiple alignments that were manually refined in GeneDoc v.2.7.0 [2].

Secondary structure was predicted by Psi-pred 2.5 [3] at the ALI2D server (http://toolkit.tuebingen.mpg.de/ali2d). Transmembrane helices were predicted with the TMHMM server v.2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/). TopPred 0.01 [4] and TMpred [5] programs were also used to make predictions of the transmembrane regions and orientations. The three different approaches gave similar results. Hydrophobic moments of amphipathic helices were predicted by MPex [6] and HeliQuest [7]. Both approaches yielded similar results.

II.2.2.2 Structural models

To generate the structural models of the antiporter-like proteins from *E. coli* (HyfB and HycC), *P. furiosus* (MbhH+MbhI and MbxH) and *M. mazei* (EchA and EhrA) we used their amino acid sequences retrieved from Uniprot. The structural models were predicted by Phyre2 server (Protein Homology/analogy Recognition Engine V 2.0) [8], which used as template the crystallographic structure of NuoL from *E. coli* (PDB code: 3RKO:L). After sequence alignment by HHpred [9] and using the same structure of NuoL as template, structural models were also predicted by MODELLER [10], providing the same result. Protein homology structure images were generated using PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.

II.2.3 Results

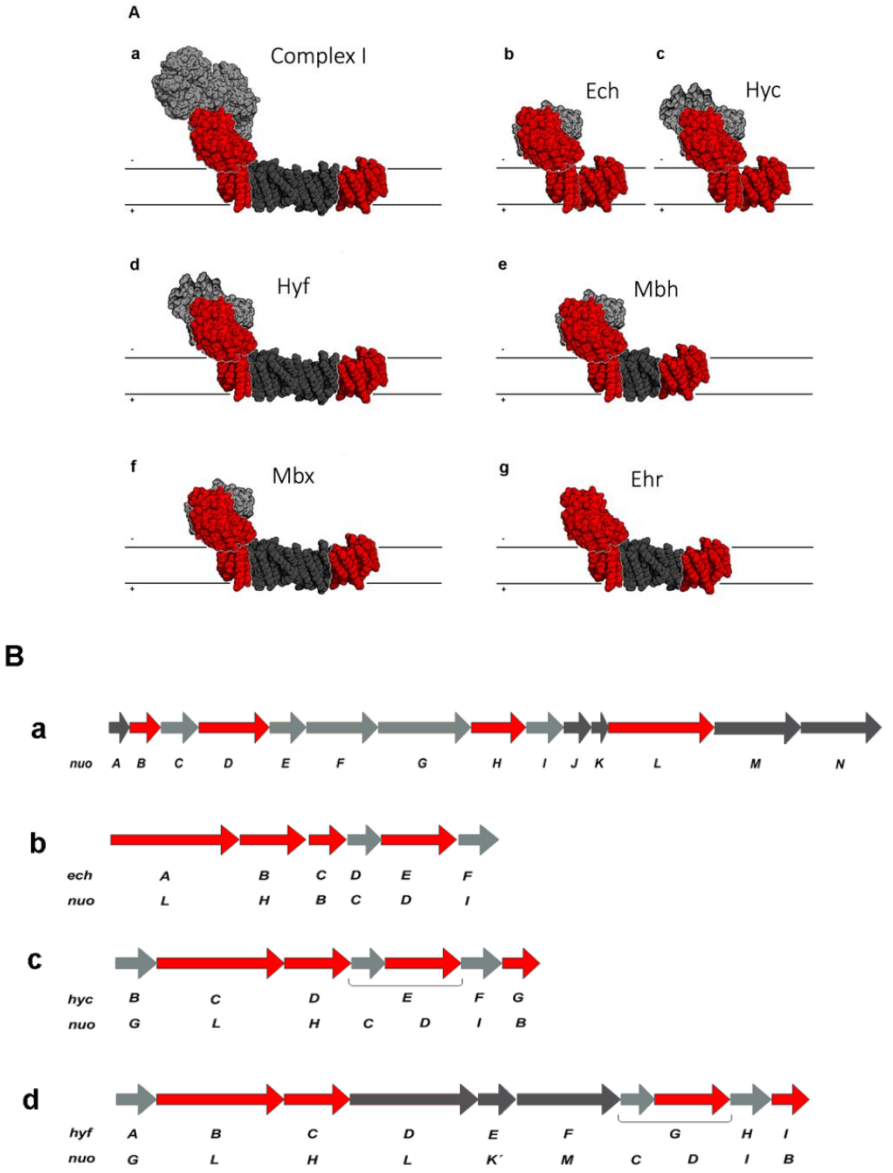
II.2.3.1 Sequence analyses reveal two types of Na^+/H^+ antiporter-like subunits, with or without a C-terminal extension

As representatives of the complex I, group 4 [NiFe] hydrogenases and Ehr family we have chosen the complexes schematized in Figure II.2.1 (panel A), whose coding gene clusters are drawn in Figure II.2.1 (panel B). To represent complex I we elected that from *E. coli* (Figure II.2.1, panels A-a and B-a), which is the most characterized and whose X-ray crystal structure has been determined [11]. The choice of the representatives of [NiFe] membrane-bound hydrogenases also took into account the available genetic and

biochemical data. The simplest hydrogenase selected was that from *M. mazei* (Figure II.2.1, panel A-b and B-b), which has already been isolated and functionally investigated [12]. The so-called Hyc hydrogenase from *E. coli* (Figure II.2.1, panel A-c and B-c) has been characterized as a formate hydrogen lyase 1 (FHL-1). This complex has, besides the subunits composing Ech, an iron-sulfur subunit homologous to the N-terminal of NuoG and a formate dehydrogenase subunit (Fdh-F), that receives electrons from formate [13]. *E. coli* contains another group 4 hydrogenase (Hyf) encoded by the *hyf* gene cluster (Figure II.2.1, panel A-d and B-d). This gene cluster is similar to *hyc*, but includes three additional genes [14]. Two of these encode homologues of the antiporter-like subunits from complex I and another (*hyfE*) encodes a protein, whose C-terminus is homologous to NuoK [14, 15]. The electron donor to this hydrogenase is still unknown. The membrane-bound hydrogenase from *P. furiosus,* Mbh, has been intensively studied and is predicted to be composed of 13 subunits, having 4 to 6 subunits homologous to complex I [16] (Figure II.2.1, panel A-e and B-e). Interestingly, we observed that in the *mbh* cluster the gene following that encoding MbhH, *mbhI*, codes for a hypothetical protein, which is homologous to the C-terminus of NuoL. In this way we considered the antiporter-like subunit of Mbh the product of the *mbhH* + *mbhI* genes. *P. furiosus* also contains a membrane complex similar to membrane-bound [NiFe] hydrogenases, but which does not have the motifs for the binding of the [NiFe] centre, like Ehrs. The complex, named Mbx (Figure II.2.1, panel A-f and B-f) has been genetically characterized and was suggested initially to

function as a ferredoxin:$NADP^+$ oxidoreductase [16] and later as a ferredoxin:sulfur oxidorecutase [17].

Although no biochemical data is available, we also included in our list of representatives of the family, the Ehr complex from *M. mazei* (Figure II.2.1, panel A-g and B-g), because this is one of the simplest hydrogenase-like complexes, which does not contain the [NiFe] binding motifs [18].



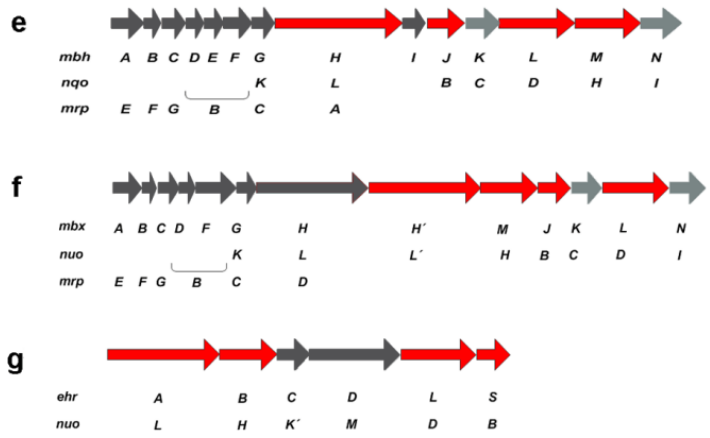
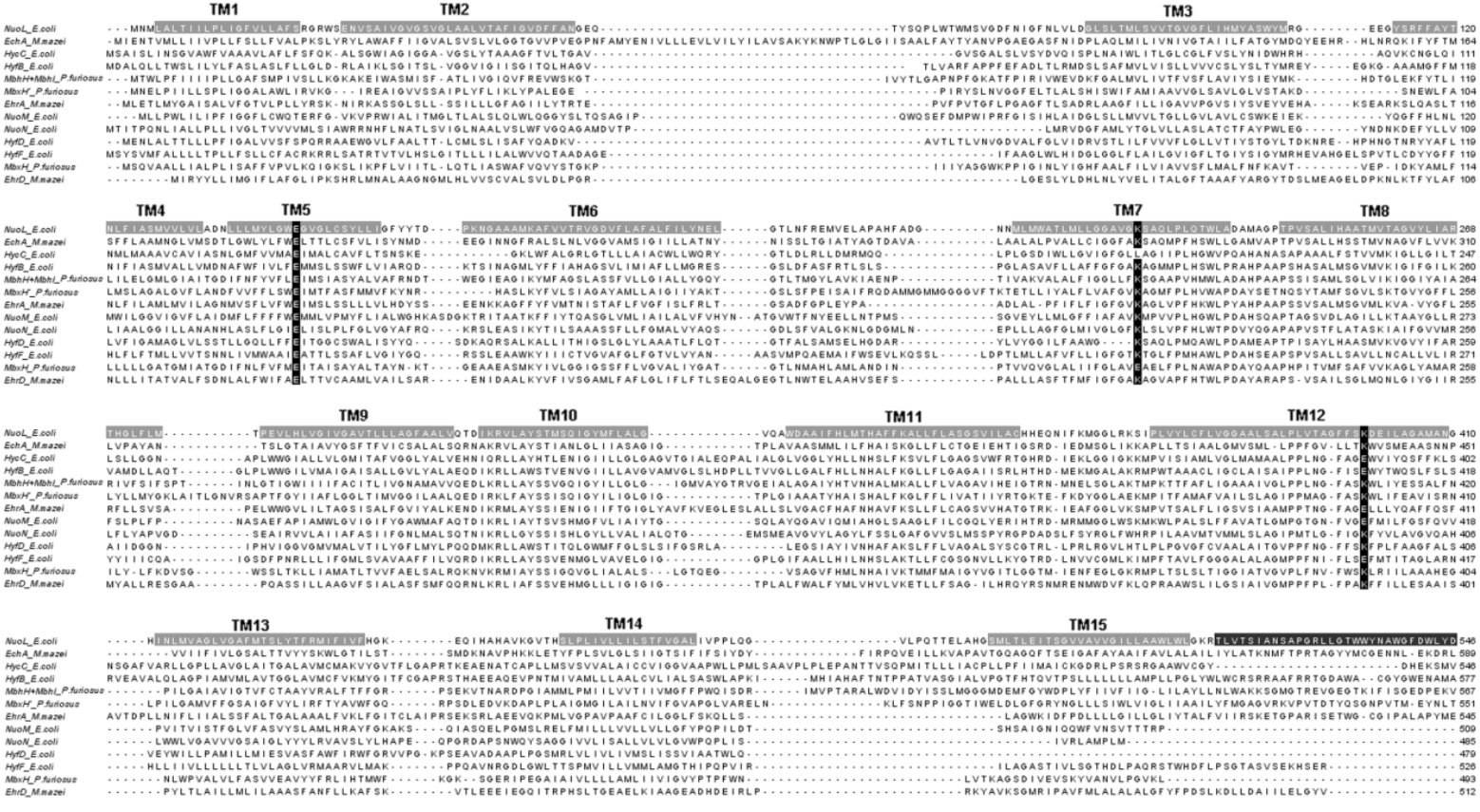


Figure II.2.1: A. Schematic representation of the protein structure of the chosen representatives of the complex I, group 4 [NiFe] hydrogenases and energy-converting hydrogenase related complexes (Ehr) family. The common subunits are coloured in red. We nominated this common denominator as the Universal Adaptor. The representatives are: complex I from *E. coli* (a), Ech from *M. mazei* (b), Hyc (c) and Hyf (d) from *E. coli*, Mbh (e) and Mbx (f) from *P. furiosus* and Ehr from *M. mazei* (g). B. Schematic representation of the gene cluster organizations coding for the chosen representatives of the complex I, group 4 [NiFe] hydrogenases and energy-converting hydrogenase related complexes (Ehr) family. Complex I from *E. coli* (a), Ech from *M. mazei* (b), Hyc (c) and Hyf (d) from *E. coli*, Mbh (e) and Mbx (f) from *P. furiosus* and Ehr from *M. mazei* (g).

We aligned, using clustalX, the sequences of all Na^+/H^+ antiporter-like subunits of the chosen representatives of the complex I, group 4 [NiFe] hydrogenases and Ehr family (Figure II.2.2). The alignment showed immediately that the sequences could be arranged in two groups according to their average sizes, having one group sequences with an approximately 100 amino acid residues longer C-terminus. The obtained similarity and identity values (Table II.2.1) do not allow to unequivocally establishing proteins such as the antiporter-like subunits from hydrogenases or Ehr as homologous of NuoL, NuoM or NuoN. Nevertheless, we can consider that the


TM1
TM2
TM3
TM4
TM5
TM6
TM7
TM8
TM9
TM10
TM11
TM12
TM13
TM14
TM15
NuoL_E.coli
EchA_M.mazei
HycC_E.coli
HyfB_E.coli
MbhH+MbhI_P.furiosus
MbxH'_P.furiosus
EhrA_M.mazei
NuoM_E.coli
NuoN_E.coli
HyfD_E.coli
HyfF_E.coli
MbxH_P.furiosus
EhrD_M.mazei

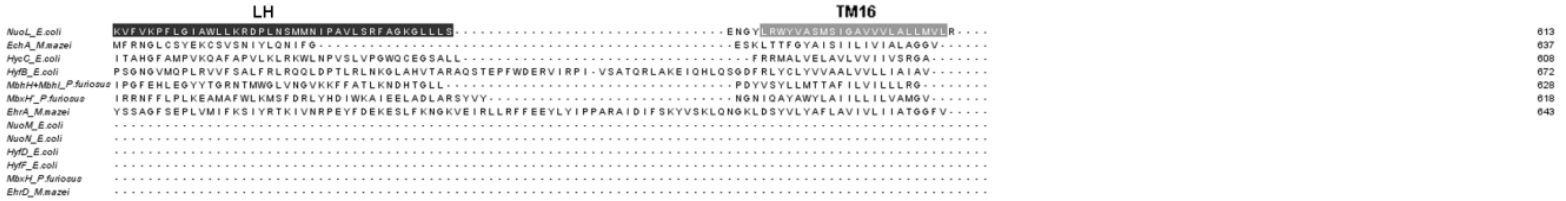


Figure II.2.2: Alignment of the amino acid sequences from the Na^+/H^+ antiporter-like subunits present in the chosen representatives of complex I, group 4 membrane-bound [NiFe] hydrogenases and Ehr family. The alignment was performed using clustalX. Glutamate residues corresponding to that at the middle of TM helix 5 of NuoL (Glu 144 – NuoL) and lysine residues corresponding to those at the middle of the discontinuous helices TM7 and TM12 of NuoL (Lys229, Lys 399 - NuoL), thought to be important in charge translocation, are highlighted in black background. The transmembrane α-helices observed in the structure of NuoL are shaded in light gray as well as the long amphipathic α-helix is shaded in medium gray.

Table II.2.1: Statistics of amino acid sequence alignment from the Na^+/H^+ antiporter-like subunits present in the chosen representatives of complex I, group 4 membrane-bound [NiFe] hydrogenases and Ehr family.

IDENTITY (%)

	NuoL	EchA	HycC	HyfB	MbhH+MbhI	MbxH´	EhrA	NuoM	NuoN	HyfD	HyfF	MbxH	EhrD
NuoL	-	19	15	17	19	16	17	15	11	21	18	16	15
EchA	-	-	14	14	17	15	15	11	11	15	15	15	16
HycC	-	-	-	30	18	18	23	13	12	15	16	14	13
HyfB	-	-	-	-	21	17	30	14	12	16	16	14	16
MbhH+MbhI	-	-	-	-	-	19	22	17	15	17	19	23	17
MbxH´	-	-	-	-	-	-	18	17	14	13	18	15	15
EhrA	-	-	-	-	-	-	-	15	12	15	18	16	17
NuoM	-	-	-	-	-	-	-	-	16	15	19	16	17
NuoN	-	-	-	-	-	-	-	-	-	15	18	19	16
HyfD	-	-	-	-	-	-	-	-	-	-	18	19	13
HyfF	-	-	-	-	-	-	-	-	-	-	-	20	24
MbxH	-	-	-	-	-	-	-	-	-	-	-	-	19
EhrD	-	-	-	-	-	-	-	-	-	-	-	-	-

SIMILARITY (%)

	NuoL	EchA	HycC	HyfB	MbhH+MbhI	MbxH´	EhrA	NuoM	NuoN	HyfD	HyfF	MbxH	EhrD
NuoL	-	36	34	36	38	37	34	32	30	37	36	33	32
EchA	-	-	31	32	37	33	32	30	26	32	32	31	32
HycC	-	-	-	51	38	37	42	28	28	34	32	32	30
HyfB	-	-	-	-	41	38	50	30	28	34	34	33	32
MbhH+MbhI	-	-	-	-	-	39	41	33	31	35	35	41	33
MbxH´	-	-	-	-	-	-	37	34	29	30	35	33	33
EhrA	-	-	-	-	-	-	-	28	29	31	33	34	31
NuoM	-	-	-	-	-	-	-	-	34	37	39	38	35
NuoN	-	-	-	-	-	-	-	-	-	36	35	40	36
HyfD	-	-	-	-	-	-	-	-	-	-	37	43	33
HyfF	-	-	-	-	-	-	-	-	-	-	-	43	44
MbxH	-	-	-	-	-	-	-	-	-	-	-	-	41
EhrD	-	-	-	-	-	-	-	-	-	-	-	-	-

antiporter-like subunit common to all representative complexes, and thus part of the universal adaptor, has a long sequence as NuoL.

Moreover, we also aligned all known sequences from EchA, HycC, HyfB, MbhH+MbhI, MbxH' and EhrA (not shown) and observed that all of them have a C-terminal extension. This observation indicated that the conclusion of the analyses performed for each representative may be generalized to the respective group of proteins.

II.2.3.2 Hydropathy profiles and secondary structure predictions of Na^+/H^+ antiporter-like subunits with C-terminal extension show the presence of a significant non transmembrane stretch between the two transmembrane regions at the C-terminus

In order to investigate whether the proteins with longer sequences have similar properties, specifically at their C-terminal extensions, we have calculated their hydropathy profiles and predicted their secondary structures (Figure II.2.3 and Supplementary Figure II.2.1). We used different algorithms for hydropathy profiling (Figure II.2.3). All proteins are predicted to have between 15 and 18 transmembrane helices, but most interesting is the observation that in all cases there is a significant non-transmembrane stretch, composed of 46 to 102 amino acid residues (Table II.2.2), between the two transmembrane helices closer to the C-terminus.

This stretch in NuoL, containing 77 amino acid residues, corresponds to the long amphipathic α-helix, i.e. one side of the helix contains mainly hydrophilic residues and the other side contains mainly hydrophobic residues. The amino acid sequence of amphipathic α-helix alternates between hydrophilic and hydrophobic residues every 3 to 4 residues, since the α-helix makes a turn for every 3.6 residues. The presence of such a helix in the other members of the family is supported by their secondary structure predictions (Figure II.2.3). These indicate that the representative proteins may have several transmembrane helices, but most relevant is the prediction of an α-helical structure for the non-transmembrane stretch present between the two transmembrane regions closer to the C-terminus. Moreover, we specifically analyzed the hydropathy of these long α-helices and we

observed an amphipathic character for all representatives as that observed for NuoL. We can conclude that the analyzed antiporter-like subunits, with extended C-terminus, present similar hydropathy profiles and secondary structure predictions to that of NuoL.

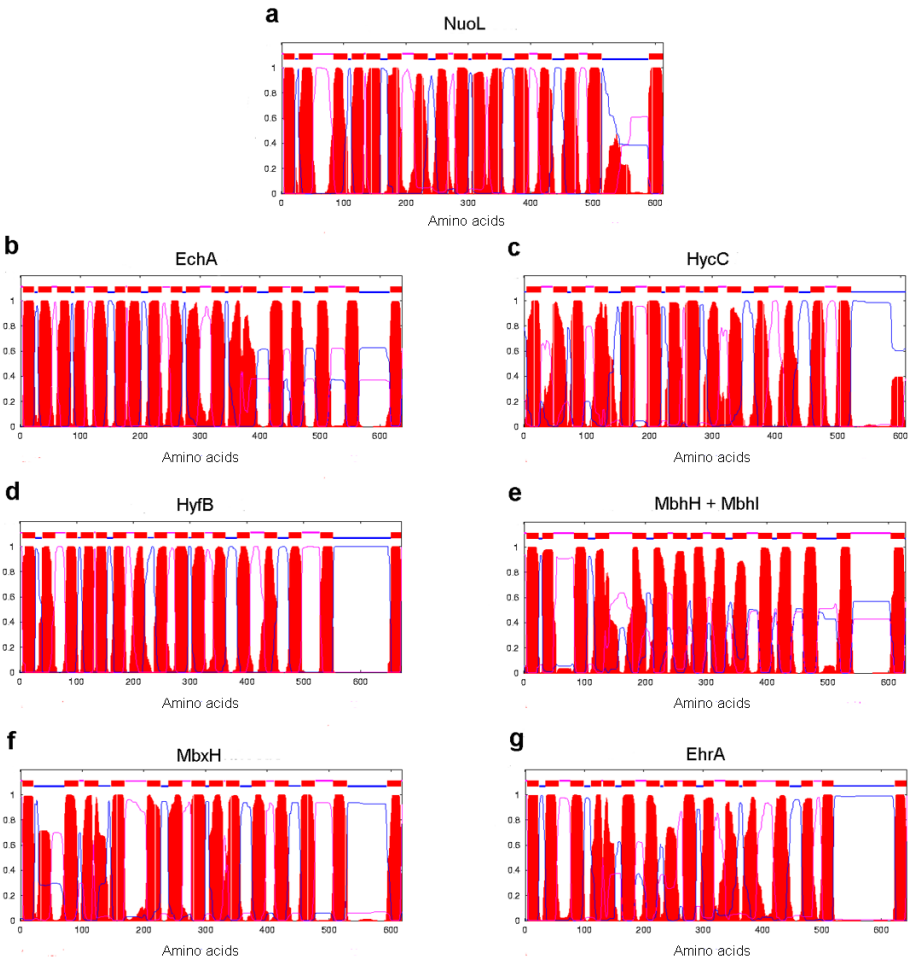


Figure II.2.3: Hydropathy profiles of the Na^+/H^+ antiporter-like subunits with extended C-terminus present in the chosen representatives of complex I, group 4 membrane-bound [NiFe] hydrogenases and Ehr family. The profiles were obtained at http://www.cbs.dtu.dk/services/TMHMM-2.0/. Other algorithms were also used to calculate hydropathy profiles yielding similar results. NuoL from *E. coli* (a), EchA from *M. mazei* (b), HycC (c) and HyfB (d) from *E. coli*, MbhH+MbhI (e) and MbxH´ (f) from *P. furiosus* and EhrA from *M. mazei* (g).

Table II.2.2: Characteristics of the long amphipathic α-helix predicted to be present in Na^+/H^+ antiporter-like subunits with C-terminal extension.

	Long α-helix	
Protein	amino acid	Length (Å)[a]
NuoL	77	115.5
EchA	46	69
HycC	51	76.5
HyfB	101	151.5
MbhH+MbhI	66	99
MbxH'	72	108
EhrA	102	153

[a] The α-helix length was estimated based on the number of amino acid residues present between the two transmembrane regions at the C-terminus of each sequence and considering that one α-helix turn contains 3.6 residues corresponding to an average length of 5.4 Å.

II.2.3.3 A long amphipathic α-helix is present in all Na^+/H^+ antiporter-like subunits with C-terminal extension

A possible existence of a HL in the common antiporter-like subunit present in the representatives of complex I, hydrogenases and Ehr is consistent with the respective sequence lengths, hydropathy profiles and secondary structure predictions. In order to further investigate this possibility we have performed homology structural models for all the representatives (Figure II.2.4). Table II.2.3 summarizes the properties of the obtained models. All models present the first 15 TMs of NuoL. EchA has two stretches of 36 and 15 amino acid residues, between NuoL's TM2 and TM3 and TM14 and TM15, respectively, which were not modeled. The presence of these extra amino acid segments seems not to alter the two inverted repeats

arrangement present in all antiporter-like subunits. Although having different lengths, HL was also modelled in all cases. With the exception for EchA and HycC, all the other models showed the presence of a 16th transmembrane helix at the C-terminus. The last 41 amino acid residues from EchA could not be modelled, but its secondary structure prediction suggests that these constitute an α-helix. Thus the presence of a 16th transmembrane helix at the C-terminus of EchA is possible. A similar situation occurs for HycC, although in this case the presence of the C-terminal α-helix is not so clear. We have performed a rough estimation of the length of the HL, based on the number of amino acid residues present between the two transmembrane regions at the C-terminus of each sequence and considering that one α-helix turn contains 3.6 residues corresponding to an average length of 5.4 Å (Table II.2.2). The estimated length for HL of NuoL, 115.5 Å, is in accordance to that obtained by the crystal structure, 110 Å [11]. The chosen representative complexes contain different numbers of antiporter-like subunits. Ech, Hyc and Mbh have one antiporter-like subunit, Ehr and Mbx may have two and complex I and possibly Hyf contain three. According to the estimations of the length of HLs, EchA and HycC are expected to have the shortest HL and HyfB contains one of the longest. Nevertheless, a correlation between the estimated length of HL and the number of antiporter-like subunits present in each complex is not clear.

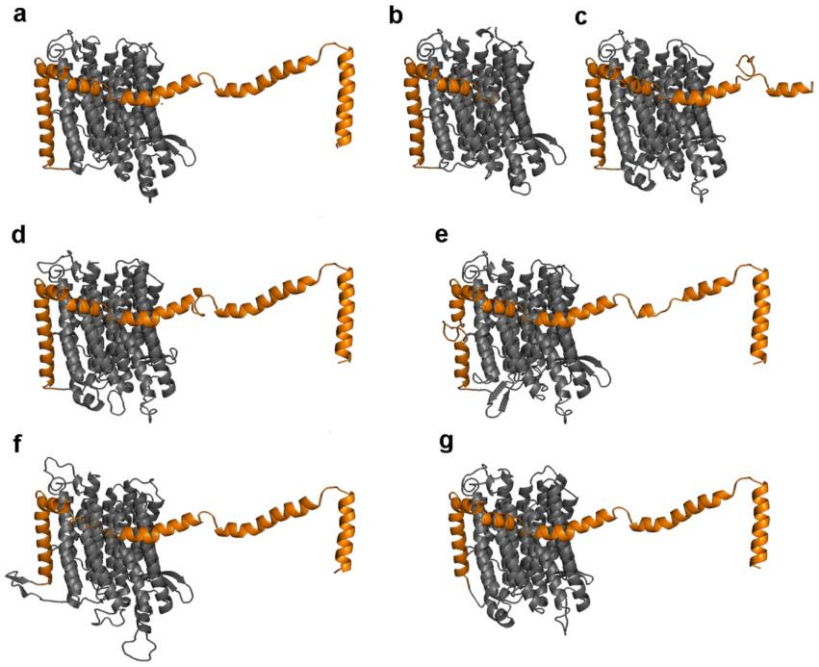


Figure II.2.4: Structure of NuoL Na^+/H^+ antiporter-like subunit from *E. coli* and homology models of the Na^+/H^+ antiporter-like subunits with extended C-terminus present in the chosen representatives of group 4 [NiFe] membrane-bound hydrogenases and Ehr family. The represented models were obtained from Phyre[2]. Side view, cytoplasmic side up. NuoL from *E. coli* (a), EchA from *M. mazei* (b), HycC (c) and HyfB (d) from *E. coli*, MbhH+MbhI (e) and MbxH´ (f) from *P. furiosus* and EhrA from *M. mazei* (g). The proteins are coloured in gray with their C-termini, including two transmembrane helices connected by a long amphipathic α-helix highlighted in orange.

II.2.3.4 Additional conserved elements among all Na^+/H^+ antiporter-like subunits with C-terminal extension

The presence of additional elements connecting the antiporter-like subunits was shown by the crystal structure of complex I. These are β-hairpins between TM helices 2 and 3 and localized close to the opposite side of the membrane in relation to that of HL. The β-hairpin from one subunit contacts its neighbor by a stretch of a C-terminal amphipathic helix. These structural arrangements were suggested to

contribute to the stability of the complex [11]. The crystal structure of the NhaA Na^+/H^+ antiporter from *E. coli* presents similar β-hairpins [19], which were suggested by cryo-EM to be at the dimer interface [20]. A NhaA mutant devoid of the β-hairpins is functional but the dimeric form was not observed [21].

We also observed that all antiporter-like subunits contain a lysine or glutamate residue at the middle of each discontinuous helices (TM7 and TM12) and at TM helix 5 (Figure II.2.2). These residues were suggested to be important for charge translocation [11]. Interestingly HycC from *E. coli* seems to be an exception since at the middle of its TM7 a leucine residue is present (Figure II.2.2). Such substitution may question the role of lysine/glutamate residues in the middle of those helices or may simply reflect an adaptive event of the organisms that contain such proteins.

Table II.2.3: Properties of the obtained structural models of the Na^+/H^+ antiporter-like subunits with C-terminal extension.

Protein		Homology model					Not modeled	
Name	aa	Template	RMSD[b]	aa	TM	%[c]	aa	aa at C-terminus[d]
EchA	638	3KOO:L	1.05	534	15	84	104	41
HycC	608	3KOO:L	1.49	546	15	90	62	24
HyfB	672	3KOO:L	1.89	596	16	89	76	21
MbhH+MbhI	628[a]	3KOO:L	1.48	600	16	96	28	2
MbxH'	618	3KOO:L	1.29	581	16	94	37	3
EhrA	643	3KOO:L	1.90	587	16	91	56	24

[a] MbhH (510 aa) + MbhI (118 aa)
[b] RMSD of backbone
[c] Amino acid sequence modelled
[d] Amino acid residues after TM15

II.2.4 Discussion

We have investigated the nature of the antiporter-like subunit common to the representative members of the complex I, hydrogenases and Ehr family. We have observed that this antiporter-like subunit contains a long amphipathic α-helix, as NuoL. Interestingly, Efremov *et al* have also suggested the presence of the HL in Hyf and Mrp complexes [22]. In this way we consider the antiporter-like subunit present in the universal adaptor as a homologue of NuoL, even though it is not unequivocally supported by sequence similarities and identities (Table II.2.1). The presence of the HL in all analyzed complexes suggests that it is also a common denominator that has been conserved. The conservation in complex I, hydrogenases and Ehr family reflects the key role of HL in the energy transduction mechanism by these complexes.

The role of HL is one of the most discussed topics in the current study of complex I. Based on the first crystallographic data the HL was suggested to have a functional role in the energy coupling mechanism of complex I. It would work like a piston in order to allow transmission between the base of the peripheral arm, where the quinone is reduced, and the membrane antiporter-like subunits, in which ion translocation is suggested to occur [11].

The available mutagenesis studies of HL produced contradictory results [23-25]. Some suggested that HL does not intervene in ion translocation and thus its function should be confined to a structural role, necessary to stabilize the membrane arm [24]. However, *E. coli* complex I mutated in an aspartate residue of the HL,

which is close to NuoM subunit, presented approximately half of the H^+/2e stoichiometry, which was interpreted as the transmission between HL and NuoL and NuoM being hampered [23].

Particularly, the HL was proposed to work as a long helical transmission element [25]. Based on the observations that complex I from *Yarrowia lipolytica,* missing two of the antiporter-like subunits (and the HL), was still catalytically active and capable of proton translocation, although at half the stoichiometry of the complete enzyme, Drose *et al* suggested that complex I has two distinct ion pump modules connected in series by the HL, which would work as a transmission component [25].

Most interesting the only antiporter-like subunit present in the simplest [NiFe] hydrogenase, EchA, contains a HL, which transverses itself. In this case a stabilizing role for HL, although not excludable, is unlikely. The compactness of the 14 core TM helices from EchA is predicted to be similar to those of other membrane proteins, including other antiporters, such as NhaA, that do not have a HL. In this way we support the suggestion that HL should have a functional role, possibly as a transmission element, in agreement with what was proposed by others [23, 25].

The presence of the β-hairpins and the lysine/glutamate residues at the middle of the discontinuous TM7 and TM12 and at TM5 in addition to the presence of the HL, reinforces the idea that antiporter-like subunits present in the members of the complex I, hydrogenases and Ehr family operate in similar ways.

Conformational changes must occur in complex I to couple the catalytic reaction to charge translocation, whether NuoL, NuoM and NuoN work as Na^+/H^+ antiporters [26-30] or as H^+ channels. In this way, element(s) responsible for such conformational changes, *i.e.* which may work as transmission devices have to be present in the different complexes. The recognition of common structural and functional denominators in the universal adaptor is a step forward in the identification of such elements and consequently in the investigation of the mechanism of complex I.

II.2.5 Acknowledgments

We specially thank Afonso S. Duarte for discussions and critical reading of the manuscript. Miguel Teixeira and Inês C. Pereira are also acknowledged for critical reading of the manuscript. A.P.B. is recipient of a fellowship from Fundação para a Ciência e a Tecnologia (SFRH/BPD/80741/2011). The project and B.C.M. fellowship were funded by Fundação para a Ciência e a Tecnologia (PTDC/QUI-BIQ/100302/2008 to M.M.P.). The work was supported by Fundação para a Ciência e a Tecnologia through grant # PEst-OE/EQB/LA0004/2011.

II.2.6 References

1. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D.,

Gibson, T. J. & Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics.* 23, 2947-8.

2. Nicholas, K. B. & Nicholas, H. B. J. (1997) GeneDoc: a tool for editing and annotating multiple sequence alignments.

3. Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices, *J Mol Biol.* 292, 195-202.

4. Claros, M. G. & von Heijne, G. (1994) TopPred II: an improved software for membrane protein structure predictions, *Comput Appl Biosci.* 10, 685-6.

5. Hofmann, K. & Stoffel, W. (1993) TMbase - A database of membrane spanning proteins segments, *Biol Chem Hoppe-Seyler.* 374.

6. Snider, C., Jayasinghe, S., Hristova, K. & White, S. H. (2009) MPEx: a tool for exploring membrane proteins, *Protein Sci.* 18, 2624-8.

7. Gautier, R., Douguet, D., Antonny, B. & Drin, G. (2008) HELIQUEST: a web server to screen sequences with specific alpha-helical properties, *Bioinformatics.* 24, 2101-2.

8. Kelley, L. A. & Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nat Protoc.* 4, 363-71.

9. Soding, J. (2005) Protein homology detection by HMM-HMM comparison, *Bioinformatics.* 21, 951-60.

10. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. & Karplus, M. (1995) Evaluation of comparative protein modeling by MODELLER, *Proteins.* 23, 318-26.

11. Efremov, R. G. & Sazanov, L. A. (2011) Structure of the membrane domain of respiratory complex I, *Nature.* 476, 414-20.

12. Welte, C., Kratzer, C. & Deppenmeier, U. (2010) Involvement of Ech hydrogenase in energy conservation of *Methanosarcina mazei*, *FEBS J.* 277, 3396-403.

13. Hedderich, R. & Forzi, L. (2005) Energy-converting [NiFe] hydrogenases: more than just H2 activation, *J Mol Microbiol Biotechnol.* 10, 92-104.

14. Andrews, S. C., Berks, B. C., McClay, J., Ambler, A., Quail, M. A., Golby, P. & Guest, J. R. (1997) A 12-cistron *Escherichia coli* operon (hyf) encoding a putative proton-translocating formate hydrogenlyase system, *Microbiology.* 143 (Pt 11), 3633-47.

15. Efremov, R. G. & Sazanov, L. A. (2012) The coupling mechanism of respiratory complex I - A structural and evolutionary perspective, *Biochim Biophys Acta.* 1817, 1785-95.

16. Bridger, S. L., Clarkson, S. M., Stirrett, K., DeBarry, M. B., Lipscomb, G. L., Schut, G. J., Westpheling, J., Scott, R. A. & Adams, M. W. (2011) Deletion strains reveal metabolic roles for key elemental sulfur-responsive proteins in *Pyrococcus furiosus*, *J Bacteriol.* 193, 6498-504.

17. Schut, G. J., Zadvornyy, O., Wu, C. H., Peters, J. W., Boyd, E. S. & Adams, M. W. (2016) The role of geochemistry and energetics in the evolution of modern respiratory complexes from a proton-reducing ancestor, *Biochim Biophys Acta*.

18. Marreiros, B. C., Batista, A. P., Duarte, A. M. & Pereira, M. M. (2012) A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases, *Biochim Biophys Acta*.

19. Hunte, C., Screpanti, E., Venturi, M., Rimon, A., Padan, E. & Michel, H. (2005) Structure of a Na+/H+ antiporter and insights into mechanism of action and regulation by pH, *Nature.* 435, 1197-202.

20. Williams, K. A. (2000) Three-dimensional structure of the ion-coupled transport protein NhaA, *Nature.* 403, 112-5.

21. Rimon, A., Tzubery, T. & Padan, E. (2007) Monomers of the NhaA Na+/H+ antiporter of *Escherichia coli* are fully functional yet dimers are beneficial under extreme stress conditions at alkaline pH in the presence of Na+ or Li+, *J Biol Chem.* 282, 26810-21.

22. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I, *Nature.* 465, 441-5.

23. Steimle, S., Willistein, M., Hegger, P., Janoschke, M., Erhardt, H. & Friedrich, T. (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I, *FEBS Lett.* 586, 699-704.

24. Belevich, G., Knuuti, J., Verkhovsky, M. I., Wikstrom, M. & Verkhovskaya, M. (2011) Probing the mechanistic role of the long alpha-helix in subunit L of respiratory Complex I from *Escherichia coli* by site-directed mutagenesis, *Mol Microbiol.* 82, 1086-95.

25. Drose, S., Krack, S., Sokolova, L., Zwicker, K., Barth, H. D., Morgner, N., Heide, H., Steger, M., Nubel, E., Zickermann, V., Kerscher, S., Brutschy, B., Radermacher, M. & Brandt, U. (2011) Functional dissection of the proton pumping modules of mitochondrial complex I, *PLoS Biol.* 9, e1001128.

26. Batista, A. P., Marreiros, B. C. & Pereira, M. M. (2011) Decoupling of the catalytic and transport activities of complex I from *Rhodothermus marinus* by sodium/proton antiporter inhibitor, *ACS Chem Biol.* 6, 477-83.

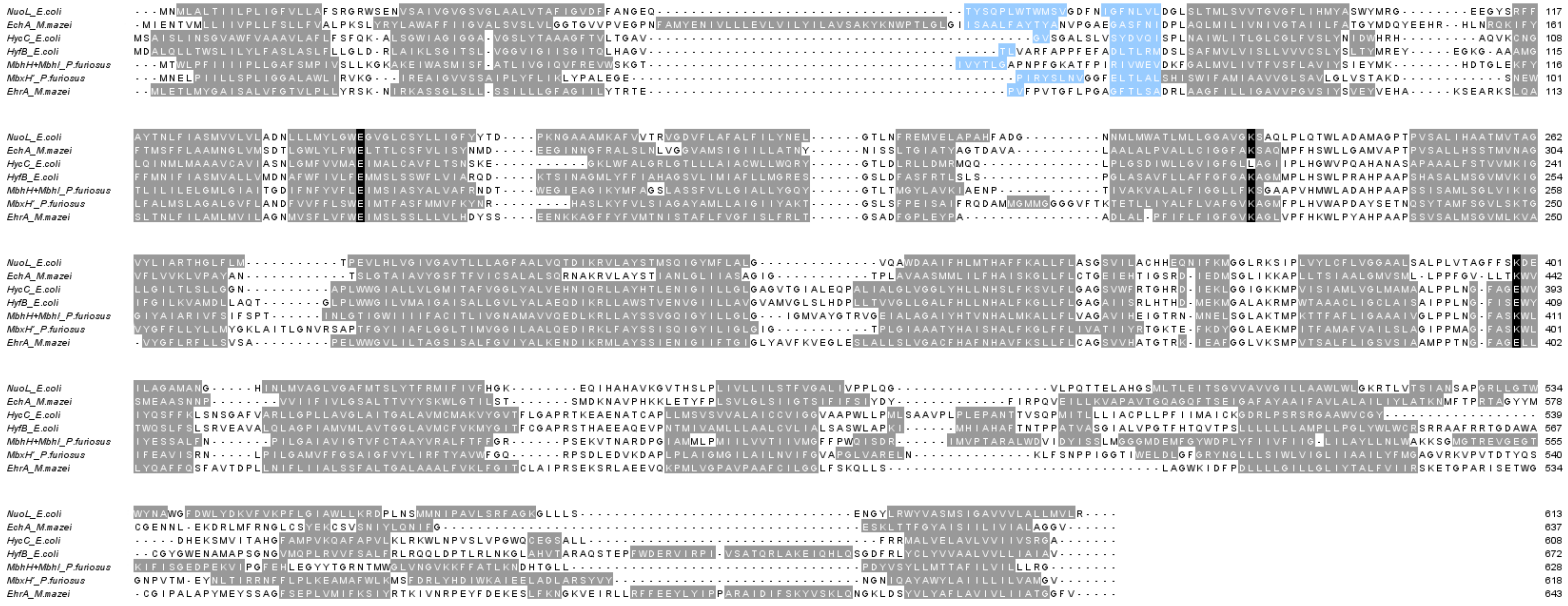
27. Batista, A. P., Fernandes, A. S., Louro, R. O., Steuber, J. & Pereira, M. M. (2010) Energy conservation by *Rhodothermus marinus* respiratory complex I, *Biochim Biophys Acta.* 1797, 509-15.

28. Stolpe, S. & Friedrich, T. (2004) The *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I) is a primary proton pump but may be capable of secondary sodium antiport, *J Biol Chem.* 279, 18377-83.

29. Steuber, J. (2003) The C-terminally truncated NuoL subunit (ND5 homologue) of the Na+-dependent complex I from *Escherichia coli* transports Na+, *J Biol Chem.* 278, 26817-22.

30. Mathiesen, C. & Hagerhall, C. (2002) Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters, *Biochim Biophys Acta.* 1556, 121-32.

II.2.7 Supplementary Material



Supplementary Figure II.2.1: Secondary structure prediction for the Na^+/H^+ antiporter-like subunits present in the chosen representatives of complex I, group 4 membrane-bound [NiFe] hydrogenases and Ehr family. The secondary structures were predicted at the Ali2D server (http://toolkit.tuebingen.mpg.de/ali2d). α-helices are shaded in gray and β-sheet are shaded in light blue. Glutamate residues corresponding to that at the middle of TM helix 5 of NuoL (Glu 144 – NuoL) and lysine residues corresponding to those at the middle of the discontinuous helices TM7 and TM12 of NuoL (Lys229, Lys 399 - NuoL), thought to be important in charge translocation, are highlighted in black background.

II.3 Na^+/H^+ antiporter activity by Complex I from *R. marinus*

This section is based on the following publication:

Batista AP, **Marreiros BC**, Pereira MM (2011) "Decoupling of catalytic and transports activities of complex I from *Rhodothermus marinus* by a sodium/proton antiporter inhibitor" ACS Chemical Biology 6 (5), 477-483. DOI:10.1021/cb100380y

Authors' contribution:

B.C.M. participated in the dual wavelength spectrophotometer experiments for the ΔΨ generation measurements and contributed for the discussion of the data.

II.3 Na^+/H^+ antiporter activity by Complex I from *R. marinus*

II.3.1 Summary

The energy transduction by complex I from *Rhodothermus marinus* was addressed by studying the influence of EIPA on the activities of this enzyme. EIPA is an inhibitor of both Na^+/H^+ antiporter and complex I NADH:quinone oxidoreductase activity. We performed studies of NADH:quinone oxidoreductase and H^+ and Na^+ translocation activities of complex I from *R. marinus* at different concentrations of EIPA, using inside-out membrane vesicles. We observed that the oxidoreductase activity and both H^+ and Na^+ transports are inhibited by EIPA. Most interestingly, the catalytic and the two transport activities showed different inhibition profiles. The transports are inhibited at concentrations of EIPA at which the catalytic activity is not affected. In this way the catalytic and transport activities were decoupled. Moreover, the inhibition of the catalytic activity was not influenced by the presence of Na^+, whereas the transport of H^+ showed different inhibition behaviors in the presence and absence of Na^+. Taken together our observations indicate that complex I from *R. marinus* performs energy transduction by two different processes: proton pumping and Na^+/H^+ antiporting. The decoupling of the catalytic and transport activities suggests the involvement of an indirect coupling mechanism, possibly through conformational changes.

II.3.2 Materials and Methods

II.3.2.1 Cell growth and membrane vesicles preparation

R. marinus PRQ 62B was grown as described previously [1], in growth medium containing 100 mM glutamate. After harvesting, cells were suspended in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4, 25 mM Na_2SO_4 (buffer A) or 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4, 50 mM choline chloride ($[Na^+]$<10 μM) (buffer B) and broken in a French Pressure cell at 19,000 psi. The membrane vesicles were obtained by ultracentrifugation of the broken cells at 200.000 g, 2 h at 4 °C followed by re-suspension in buffer A or buffer B. Integrity of vesicles was checked by the K^+/valinomycin assay using oxonol VI as a ΔΨ sensitive dye. Protein concentration was determined by the Biuret method modified for membrane proteins [2].

II .4.2.2 Protein Purification

Protein was purified as described in [3], introducing an additional chromatographic step using a Mono Q column and 20 mM Tris-HCl pH 8, 1 mM phenylmethanesulfonyl fluoride, 0.1% DDM (w/v) as buffer. The complex was eluted in a linear gradient of 0 to 100% NaCl (1 M). Protein concentration was determined by the Bicinchoninic Acid Method as described before [4].

II .4.2.3 ΔΨ detection

ΔΨ generation was detected following oxonol VI absorption (A_{628} *minus* A_{587}) at 27 °C, on an OLIS upgraded Aminco DW2 dual wavelength spectrophotometer [5]. The integrity of the vesicles was

verified by generating K^+ gradients with K^+/valinomycin in an external buffer containing 50 mM K_2SO_4 (internal [K^+]=10 mM). The assay was started adding 2 µM valinomycin. To detect the NADH-driven formation of ΔΨ, 4 mM K_2-NADH was added to vesicles in buffer A or buffer B. When referred, ethyl-isopropyl amiloride (EIPA) (0 to 100 µM), carbonyl cyanide m-chlorophenyl hydrazone (CCCP) (10 and 100 µM), rotenone (10 µM) and KCN (5 mM) were added prior to the addition of NADH.

II.3.2.4 Determination of the internal volume of membrane vesicles

The internal volume of the vesicles was determined by EPR spectroscopy, using 2,2,6,6-Tetramethyl-1-piperidinyloxy (TEMPO) oxidized with $K_3[Fe(CN)_6]$ [6]. TEMPO was quenched with 100 mM of potassium chromium(III) oxalate, in the external medium. EPR measurements were performed in a Bruker EMX spectrometer at room temperature, with a microwave frequency of 9.39 GHz, microwave power 1 mW and modulation amplitude 0.04 mT.

II.3.2.5 Activity measurements

Oxygen consumption was measured with a Clark-type oxygen electrode YSI Model 5300. The assay mixture contained vesicles in buffer A or buffer B. The reaction was started by adding K_2-NADH. When used, EIPA (0 to 800 µM), valinomycin (2 µM) or CCCP (10 µM) were added prior to the addition of substrate. NADH:quinone oxidoreductase activity was monitored on an OLIS upgraded Aminco DW2 dual wavelength spectrophotometer, at 55 °C, following the

NADH oxidation at 330 nm (ε= 5930 $M^{-1}cm^{-1}$). The reaction mixture contained 200 mM KPi pH 7.5, 0.1% DDM (w/v) and 0 ([Na^+]<10µM) or 25 mM Na_2SO_4, 50 µM 2,3-dimethyl-1,4-naphthoquinone (DMN) and 50 µM NADH. The reaction was started by the addition of complex I. When used, NaCl (0 to 200 mM) or EIPA (0 to 250 µM) were added prior to the addition of substrate. NADH:$K_3[Fe(CN)_6]$ oxidoreductase activity was monitored at 420 nm (ε= 1020 $M^{-1}cm^{-1}$). The reaction medium contained membrane vesicles or solubilized membranes in buffer A or buffer B, 250 µM $K_3[Fe(CN)_6]$ and 250 µM NADH. Solubilized membranes were obtained by stirring an aliquot of membrane vesicles with 10% of DDM (w/v) for 4 h at 4 °C.

II.3.2.6 Fluorescence spectroscopy

Fluorescence spectroscopy was performed on a Varian Cary Eclipse spectrofluorimeter. Generation of ΔpH was determined by the quenching of the fluorescence of 9-amino, 6-chloro, 2-methoxyacridine (ACMA) ($\lambda_{excitation}$=410 nm, $\lambda_{emission}$=480 nm). Vesicles were incubated aerobically for 5 min at 27 °C in buffer A or buffer B containing 1 µM of ACMA. The reaction was started by adding 100 µM of NADH. When referred, valimonycin (2 µM) and EIPA (0 to 25 µM) were added prior to the addition of NADH.

The extra-vesicle pH (pH_{out}) was monitored using the hydrophilic and membrane impermeable pH indicator, pyranine ($\lambda_{excitation}$=458 nm, $\lambda_{emission}$=508 nm). Vesicles were incubated in buffer A or buffer B containing 2 µM pyranine. The reaction was started by adding 4 mM NADH. When referred, valimonycin (2 µM), rotenone (10

µM), KCN (5 mM), DMN (200 µM) or EIPA (0 to 22.5 µM) were added prior to the addition of NADH.

II.3.2.7 ^{23}Na-NMR spectroscopy

NMR spectra were recorded on a Bruker Avance II 500 MHz spectrometer at 27 °C, operating at 132 MHz for ^{23}Na [7]. Methylenephosphonate (4.5 mM) was used as a shift reagent for the Na^+ signal of the suspension medium. A capillary tube containing the shift reagent dysprosium (III) tripolyphosphate (22 mM), was used in all experiments as external reference [8]. Spectra were recorded upon addition of 4 mM K_2-NADH to vesicles which were previously incubated with KCN (10 mM), DMN (200 µM), valinomycin (2 µM) or EIPA (0 to 20 µM).

II.3.3 Results and Discussion

We have previously shown that complex I from *R. marinus* performs three activities: 1) the catalytic activity NADH:quinone oxidoreductase, 2) H^+ translocation in the same direction of the established ΔΨ and 3) Na^+ transport to the opposite direction of 2) [7]. In this work we studied the influence of a Na^+/H^+ antiporter inhibitor, EIPA, on these activities. We use membrane vesicles from *R. marinus*, with an internal volume of 1 µL *per* mg of protein. These vesicles allowed the formation of a membrane potential using K^+ and valinomycin, which was stable for at least 20 min. This observation indicated that the *R. marinus* vesicles were tight. The orientation of

the vesicles was determined comparing the NADH:$K_3[Fe(CN)_6]$ oxidoreductase activity of vesicles before (0.54 ± 0.12 µmol $K_3[Fe(CN)_6]$ min^{-1} mg^{-1}) and after (0.55 ± 0.14 µmol $K_3[Fe(CN)_6]$ min^{-1} mg^{-1}) their solubilisation with n-Dodecyl-beta-Maltoside (DDM). The activity was, approximately, the same in both situations showing that the membrane vesicles obtained had an inside-out orientation.

II.3.3.1 Effect of EIPA on NADH-driven membrane potential (ΔΨ) generation

The addition of the substrate NADH to the vesicles originated a jump in the absorbance difference, A_{628nm}–$A_{587nm,}$ of oxonol VI, indicating the establishment of a ΔΨ (positive inside) (Figure II.3.1 and II.3.2). When the vesicles were pre-incubated with the terminal oxygen reductases inhibitor (KCN) the jump in the absorbance difference observed upon substrate addition was negligible. This finding indicates that the observed ΔΨ was created by the functioning of the respiratory chain (Figure II.3.1 panel a, trace b). A similar behavior was observed in the presence of rotenone (inhibitor of complex I) or in the presence of the protonophore CCCP (Figure II.3.1 panel a, traces c and d and Figure II.3.2). These observations are in agreement with the results obtained before [7]. Approximately the same trend of absorbance difference was detected in the presence of 25 mM Na_2SO_4 or 50 mM choline chloride ($[Na^+]$ < 10 µM) (II.3.1 panels b, c and d, traces a) indicating that sodium ions do not influence the establishment of ΔΨ.

The effect of EIPA on the ΔΨ was tested in the presence and absence of Na^+ (Figure II.3.1 panels b, c and d). Different concentrations of EIPA were added to the reaction mixture prior to the addition of NADH. For EIPA concentrations up to 20 µM, no clear effect was observed, independently of the presence of Na^+ (Figure II.3.1 panels b, c and d, traces a and b). Approximately 50 % inhibition of ΔΨ formation was observed at 60 µM EIPA for both conditions (Figure II.3.1 panel d).

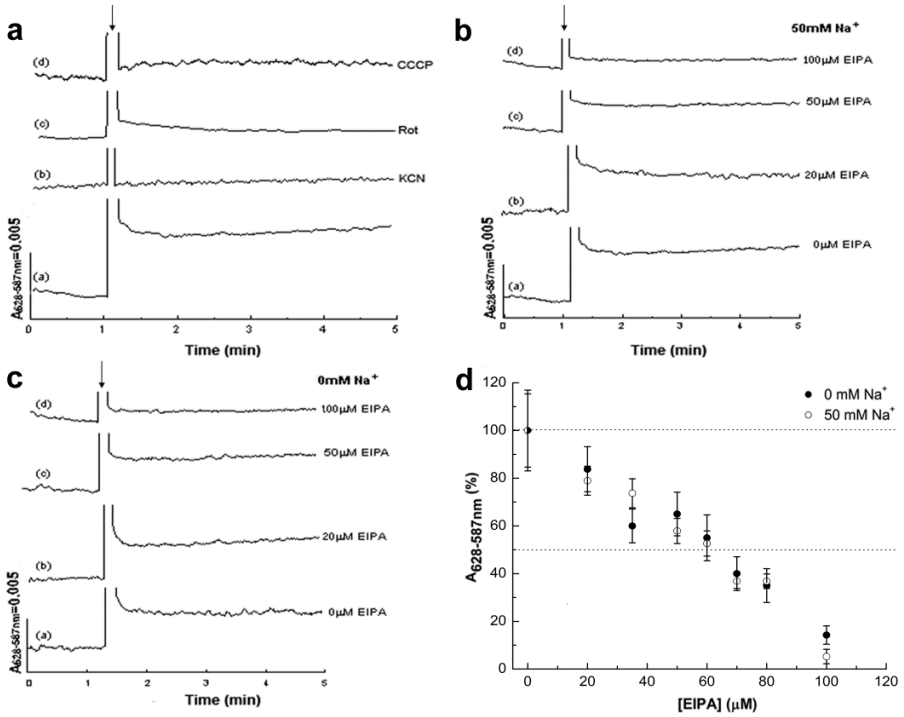


Figure II.3.1: Generation of a membrane potential (ΔΨ) by the functioning of *R. marinus* respiratory chain, monitored by the absorbance difference ($A_{628\text{-}587nm}$) of oxonol VI. Membrane vesicles were prepared in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4, 25 mM Na_2SO_4 or 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4, 50 mM choline chloride ($[Na^+]$<10µM). The reaction was started by addition of 4 mM NADH (indicated by an arrow). Panel a: Absorbance difference of oxonol VI, in the presence

of Na^+, upon the addition of NADH to vesicles (a) and pre-incubated vesicles with 5 mM KCN (b), 10 µM rotenone (Rot) (c) or 10 µM CCCP (d). Panels b, c and d: Influence of EIPA on the NADH-driven ΔΨ in the presence (panel b, open circles panel d) or the absence (panel c, filled circles panel d) of Na^+. The represented data are the average ± standard deviation of at least three independent assays (panel d). NADH-driven ΔΨ formation in the absence of EIPA was used as reference (100%).

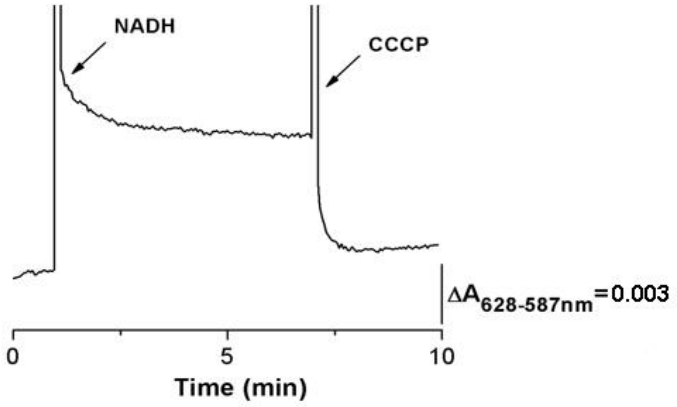


Figure II.3.2: Generation of a ΔΨ by the functioning of *R. marinus* respiratory chain, monitored by the absorbance difference ($A_{628-587nm}$) of oxonol VI. Membrane vesicles were prepared in buffer A. The reaction was started by the addition of 4 mM NADH (indicated by an arrow) and the ΔΨ was dissipated by the addition of 100 µM CCCP (also indicated by an arrow).

II.3.3.2 Effect of EIPA on the NADH:dioxygen oxidoreductase activity

R. marinus vesicles were capable of consuming oxygen upon addition of NADH, with an approximate rate of 23 nmol O_2 mg^{-1} min^{-1} at 27 °C. As previously observed [7], this activity was not dependent or stimulated by the presence of Na^+. In the presence of valinomycin and CCCP the consumption of oxygen increased (~38 nmol O_2 mg^{-1} min^{-1}), corroborating the above results that oxidation of NADH contributed to the generation of a membrane potential in tight vesicles. The NADH:oxygen oxidoreductase activity was affected by EIPA, revealing

a 10–20% inhibition in the presence of 50 µM EIPA (Figure II.3.3). The inhibition increased with the increase of the concentration of EIPA. The effect of EIPA on the NADH:dioxygen oxidoreductase activity was not affected by the presence of Na^+ and approximately 600 µM of EIPA for half maximal inhibition (IC_{50}) was estimated (Figure II.3.3).

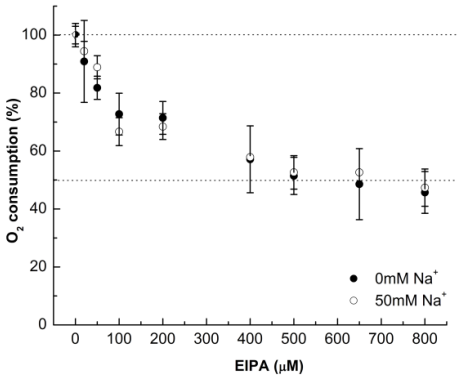


Figure II.3.3: Effect of EIPA on the NADH:O_2 oxidoreductase activity of *R. marinus* membrane vesicles. The activity was measured in the presence (open circles) or absence (filled circles) of Na^+. The reaction was started by addition of 4 mM NADH. The represented data are the average ± standard deviation of at least three independent assays. NADH:O_2 oxidoreduction in the presence of Na^+ and in the absence of EIPA was used as reference (100%).

II.3.3.3 Effect of EIPA on the NADH:DMN oxidoreductase activity

NADH:quinone oxidoreductase activity of purified complex I from *R. marinus* was determined using the menaquinone analogue, DMN, at different concentrations of Na^+ (0 to 200 mM). The specific activity was 0.8 µmol of NADH min^{-1} mg^{-1} in the absence of Na^+, a value in the range of those determined for other isolated bacterial enzymes in the presence of detergent, without the addition of lipids [9]. NADH:DMN oxidoreductase activity was not stimulated by the presence of Na^+; on the contrary, a slight decrease in activity was

observed upon increasing Na^+ concentration (Figure II.3.4 panel a). It is worth mentioning that such a decrease was not observed in the case of vesicles.

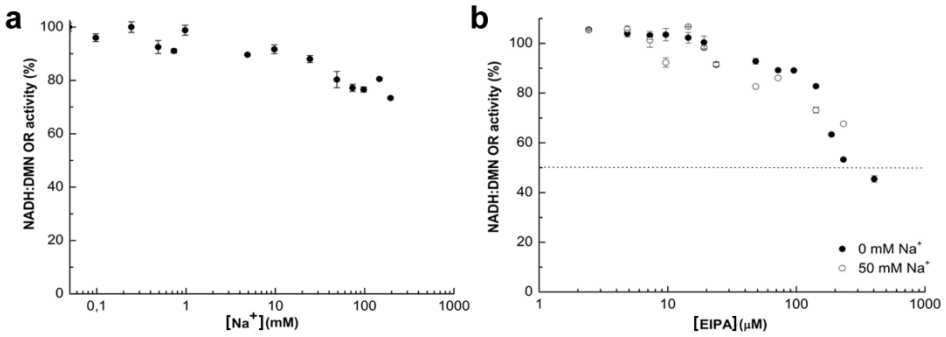


Figure II.3.4: NADH:2,3-dimethyl-1,4-naphthoquinone (DMN) oxidoreductase activity profiles of purified complex I from *R. marinus*. The assay mixture contained 24 μg of purified complex I in 200 mM KPi pH 7.5 with 0.1% n-dodecyl-β-D-maltoside, 50 μM NADH and 50 μM DMN. The reaction was started by addition of complex I and was monitored at 330 nm. The represented data are the average ± standard deviation of at least three independent assays. Panel a: Influence of the Na^+ concentration on the catalytic activity. NADH:DMN oxidoreduction in the absence of Na^+ ([Na^+]<10 μM) was used as reference (100%). Panel b: Effect of EIPA on the catalytic activity in the presence (open circles) and absence (filled circles) of Na^+. NADH:DMN oxidoreduction in the absence of EIPA was used as reference (100%).

It is well established in the literature that EIPA inhibits catalytic activity of complex I, most probably by blocking the quinone binding site. It has been observed that 100 μM of EIPA can prevent labeling of bovine complex I with a fenpyroximate analogue, a specific inhibitor of complex I which is assumed to bind at or close to a quinone binding site [10]. In the case of complex I from *R. marinus* no inhibition of NADH:quinone oxidoreductase activity was observed at least up to 20 μM EIPA and an IC_{50} for EIPA of approximately 230 μM was estimated,

irrespectively of the presence of Na^+ (Figure II.3.4 panel b and Table II.3.1). This IC_{50} value is in the range of those determined for reconstituted complex I from *E. coli* (IC_{50}=100 µM) [9], for complex I in bacterial membrane vesicles (*E. coli and P. denitrificans*) and sub-mitochondrial particles [10].

Table II.3.1: Inhibition of activities of complex I from *R. marinus*: NADH:quinone oxidoreductase activity (NADH:Q OR), proton transport (ΔpH_{out}) and sodium transport (ΔpNa^+_{out}), by EIPA.

% of inhibition

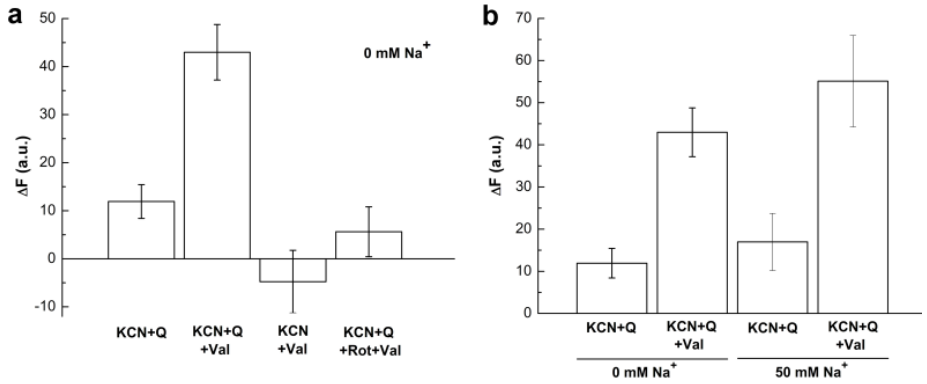
	0 mM Na^+		50 mM Na^+		
[EIPA] (µM)	NADH:Q OR	ΔpH_{out}	NADH:Q OR	ΔpH_{out}	ΔpNa^+_{out}
10	0	0	0	32	50
20	0	46	0	70-80	70-80
230	50	-[a]	50	-[a]	-[a]

[a]At 230 µM EIPA, NADH:quinone oxidoredutase activity is already 50% inhibited thus ΔpH_{out} and ΔpNa^+_{out} are also directly affected by that inhibition.

II.3.3.4 Effect of EIPA on NADH-driven external-vesicle pH (pH_{out}) change

Considering the EIPA inhibitory profile of the catalytic reaction, concentrations between 0 to 20 µM (a range in which neither the catalytic reaction nor the integrity of the membrane vesicles are affected) were chosen for the ion translocations studies. Changes on the pH_{out}, due to the activity of complex I, were measured using the pyranine fluorescence probe. Pyranine is a hydrophilic and membrane

impermeable pH indicator that responds rapidly to changes in pH in the physiological range [11]. The vesicles were pre-incubated with KCN and DMN, the inhibitor of the oxygen reductases and a menaquinone analogue, respectively, that allowed us to focus on the NADH-quinone segment of the respiratory chain from *R. marinus*. An increase in the pyranine fluorescence intensity is observed, upon addition of NADH, indicating an increase in the pH_{out} due to the functioning of complex I (Figures II.3.5 panels a and b). In the presence of valinomycin more than a two-fold increase of the pH_{out} change was observed revealing that NADH:DMN oxidoreductase activity was previously limited by ΔΨ (Figures II.3.5 panels a and b). When the vesicles were pre-incubated with KCN (without DMN) or rotenone the alkalinization upon NADH addition was considerably impaired (Figure II.3.5, panel a). The observed increase by the presence of valinomycin, was the same in the presence and absence of Na^+. However, the change in pH_{out} is more pronounced when Na^+ is present (Figure II.3.5, panel b).



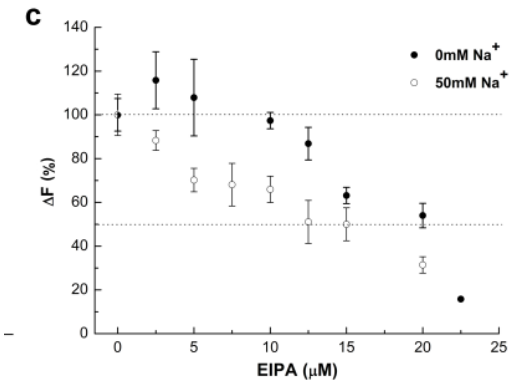


Figure II.3.5: Effect of membrane potential (ΔΨ), inhibitors of respiratory chain complexes and EIPA on NADH-driven pH_{out} variation by *R. marinus* complex I. The pyranine fluorescence was measured using a $\lambda_{excitation}$=458 nm and a $\lambda_{emission}$=508 nm. The reaction was started by addition of 4 mM NADH. The represented data are the average ± standard deviation of at least three independent assays. Panel a: Influence of ΔΨ and of inhibitors on NADH-driven pH_{out} variation. The assay mixture contained vesicles, prepared in the absence of Na^+, pre-incubated with: KCN (5 mM) and 2,3-dimethyl-1,4-naphthoquinone (DMN) (200 μM) (KCN+Q); KCN (5 mM), DMN (200 μM) and valinomycin (2 μM) (KCN+Q+Val); KCN (5 mM) and valinomycin (2 μM) (KCN+Val) and KCN (5 mM), DMN (200 μM), rotenone (10 μM) and valinomycin (2 μM) (KCN+Q+Rot+Val). Panel b: Influence of ΔΨ on NADH-driven pH_{out} variation in the presence and absence of Na^+. The assay mixture contained vesicles, prepared in the absence or presence of Na^+ (50 mM Na^+), pre-incubated with KCN (5 mM), DMN (200 μM) with or without valinomycin (Val) (2 μM). Panel c: Influence of EIPA on NADH-driven pH_{out} variation. The assay mixture contained vesicles, prepared in the absence (filled circles) or presence of Na^+ (open circles), pre-incubated with KCN (5 mM), DMN (200 μM) and valinomycin (2 μM). NADH-driven pH_{out} variation in the presence of 0 μM of EIPA was used as reference (100%). See Figure II.3.6 for examples of kinetic traces.

The influence of EIPA on the NADH-driven pH_{out} change was furthermore investigated using different concentrations of EIPA, in the presence or absence of Na^+ (Figure II.3.5 panel c and Figure II.3.6). These assays were performed in the presence of valinomycin in order to overcome the limitation of activity by ΔΨ. In the absence of Na^+, ~46% of inhibition was reached at 20 μM EIPA and no inhibition was

detected at 10 µM (filled circles in Figure II.3.4 panel c, Table II.3.1 and Supplementary Figure II.3.2, panel a). In the presence of Na^+, at 10 µM EIPA a decrease in the change of the pyranine fluorescence intensity of ~32% was observed and at 20 µM EIPA the decrease in that change was ~70% (open circles in Figure 4 panel c, Table 1 and Supplementary Figure 2 panel b). An IC_{50} value of 21±1 µM of EIPA was estimated for H^+ transport in the absence of Na^+ and 13.5±1.5 µM in the presence of this ion.

These results show different inhibitory behaviors for complex I proton transport by EIPA in the presence and absence of Na^+. In its absence, 20 µM EIPA was responsible for a decrease of ~46% in ΔpH_{out}. Since at this concentration of EIPA, the NADH:quinone oxidoreduction was not affected (Figure II.3.3 panel b and Table II.3.1), the observed decrease in ΔpH_{out} was not due to a decrease in proton consumption for quinone reduction, but originates from a decrease in proton uptake for translocation. Therefore we conclude that in those conditions EIPA inhibited H^+ translocation and promoted decoupling of the oxidoreductase and H^+ translocating activities. The decoupling of H^+ transport from the catalytic reaction was further supported by the assays performed in the presence of Na^+ in which 5 µM EIPA was sufficient to inhibit ~30% of the H^+ transport.

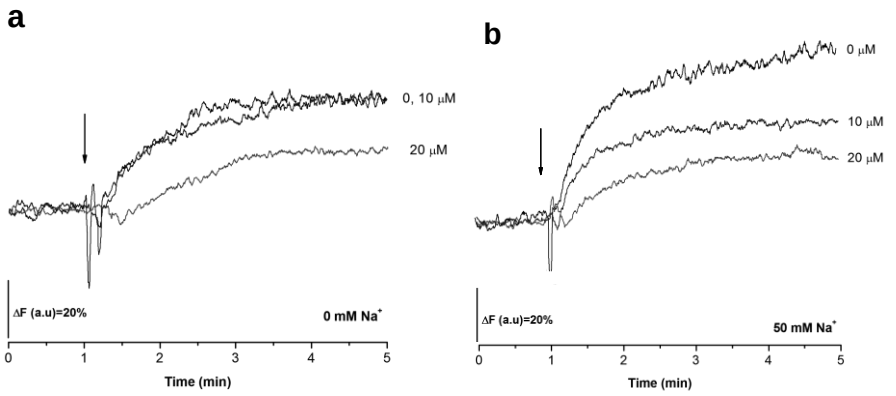


Figure II.3.6: Examples of kinetic traces of the effect of EIPA (0, 10 and 20 μM) on NADH-driven ΔpH_{out} variation by *R. marinus* complex I. The pyranine fluorescence was detected using $\lambda_{excitation}$=458 nm and $\lambda_{emission}$=508 nm. The reaction was started by the addition of 4 mM NADH (indicated by an arrow). The assay mixture contained membrane vesicles, prepared in the absence (Panel a) or presence of Na^+ (Panel b), pre-incubated with KCN (5 mM), DMN (200 μM) and valinomycin (2 μM).

II.3.3.5 Effect of EIPA on ΔpH generation

The decoupling of H^+ transport from the catalytic activity was further corroborated by the studies performed with the integral respiratory chain from *R. marinus*, in which an inhibition of the proton transport was observed at EIPA concentrations at which the catalytic activity was not affected. It was previously shown that the functioning of complex I from *R. marinus* allowed the establishment of a ΔpH, which was completely sensitive to the presence of the protonophore CCCP and of the Na^+/H^+ exchanger monensin. Although the presence of Na^+ was not a requisite for the establishment of ΔpH, it increased the ΔpH [7]. In the current study, the effect of EIPA on the NADH-driven ΔpH was monitored using ACMA in the presence and absence of Na^+ (Figure II.3.7). Since we have observed an NADH-driven ΔpH increase of 20-30% in the absence of ΔΨ, these assays were

performed in the presence of valinomycin and K^+, in order to guarantee that the establishment of ΔpH was not limited by ΔΨ. In the presence of Na^+, ~30% and ~45% decrease of the ΔpH amplitude was observed at 10 and 20 µM of EIPA, respectively. A corresponding IC_{50} value of 20.5±0.5 µM was determined (Figure II.3.7 panel a and Figure II.3.8). When Na_2SO_4 was replaced by choline chloride, only ~20% inhibition of ΔpH amplitude was observed at 20 µM EIPA. In the absence of Na^+, 10 µM EIPA had no effect on the NADH-driven ΔpH (Figure II.3.7 b and Figure II.3.8).

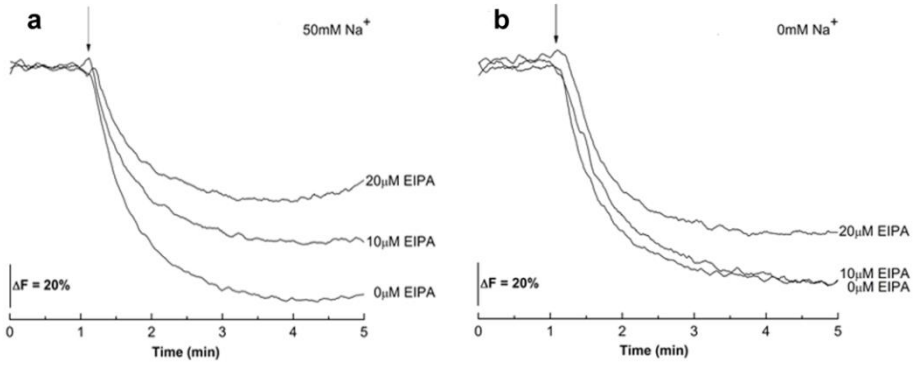


Figure II.3.7: Effect of EIPA on NADH-driven ΔpH generation by the respiratory chain of *R. marinus*. The 9-amino-6-chloro-2-methoxyacridine (ACMA) fluorescence was measured using $\lambda_{excitation}$=410 nm and $\lambda_{emission}$=480 nm. The reaction was started by the addition of 100 µM NADH (indicated by an arrow). Influence of EIPA (0, 10 and 20 µM) on NADH-driven ΔpH generation by vesicles prepared in the presence (panel a) or absence (panel b) of Na^+.

Our observations are in agreement with the results obtained by Dlasková and coworkers, who observed inhibition of H^+ transport by EIPA using glutamate and malate as substrates of rat liver mitochondria without affecting respiration [12]. An apparent IC_{50} value for EIPA inhibition of that transport, when added before the

substrates was estimated to be 27 µM. The decoupling effect was assigned to decoupling of complex I. The authors also observed no inhibition of mitochondrial respiration of the mentioned substrates until EIPA concentration reached 250–500 µM [12].

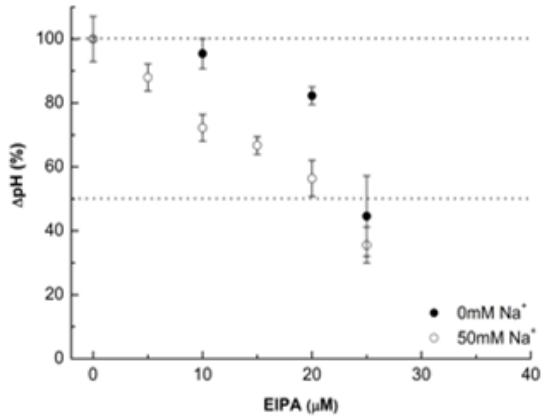


Figure II.3.8: Effect of EIPA on NADH-driven ΔpH generation by the respiratory chain of *R. marinus*. The 9-amino-6-chloro-2-methoxyacridine (ACMA) fluorescence was detected using an $\lambda_{excitation}$=410 nm and an $\lambda_{emission}$=480 nm. Influence of EIPA (0 to 25 µM) on NADH-driven ΔpH generation by vesicles prepared in the presence (open circles) or absence (filled circles) of Na^+. The represented data are the average ± standard deviation of at least three separated assays. NADH-driven ΔpH variation in the presence of 0 µM of EIPA was used as reference (100%).

II.3.3.6 Effect of EIPA on Na^+-transport

It was previously shown using ^{23}Na-NMR spectroscopy that NADH consumption was accompanied by Na^+ transport from the inside to the outside of *R. marinus* vesicles, which was completely sensitive to KCN, monensin and rotenone. This transport was actively performed since it was stimulated by the protonophore CCCP, excluding a transport by a secondary event dependent on a proton gradient or simply by a charge compensation event [7]. The effect of

EIPA on Na^+ transport by complex I from *R. marinus* was also investigated using vesicles pre-incubated with KCN and DMN. After NADH addition, an increase in the external medium of 17±1.3 nmol of Na^+ *per* mg of protein was observed (Figure II.3.9) and an increase of ~42% was observed in the presence of valinomycin (Supplementary Figure II.3.4), showing that Na^+ transport was limited by ΔΨ. In the presence of valinomycin, EIPA inhibits the external increase of Na^+ concentration with an estimated IC_{50} of 9±1 µM (Figure II.3.10).

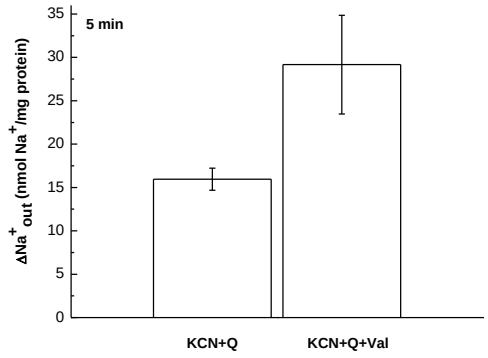


Figure II.3.9: Effect of ΔΨ on NADH-driven $\Delta[Na^+]_{out}$ by *R. marinus* complex I monitored by ^{23}Na-NMR spectroscopy. The assay mixture contained vesicles pre-incubated with KCN (5 mM) and DMN (200 µM) in the absence (KCN+Q) or presence of valinomycin (2 µM) (KCN+Q+Val). The reaction was started by the addition of 4 mM NADH. The represented data are the average ± standard deviation of at least three separated assays.

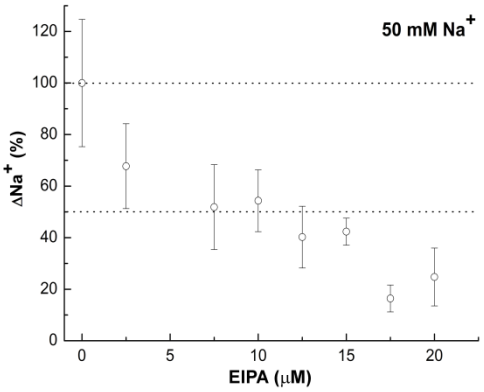


Figure II.3.10: Effect of EIPA (0 to 20 μM) on NADH-driven $\Delta[Na^+]_{out}$ by complex I from *R. marinus* monitored by ^{23}Na-NMR. The reaction was started by the addition of 4 mM NADH. The assay mixture contained vesicles pre-incubated with KCN (5 mM), DMN (200 μM) and valinomycin (2 μM). The represented data are the average ± standard deviation of at least three independent assays. NADH-driven $\Delta[Na^+]_{out}$ in the presence of 0 μM of EIPA was used as reference (100%).

Taken together, the data from the H^+ and Na^+ transport studies reveal that in the presence of Na^+, 10 μM EIPA was sufficient to inhibit both H^+ and Na^+ transports, but to different extensions (Table II.3.1). In these conditions, decreases of ~32% and ~50% were observed for NADH-driven pH_{out} and NADH-driven $[Na^+]_{out}$ changes, respectively. At 20 μM EIPA, both the pH_{out} and $[Na^+]_{out}$ changes decreased to ~70-80%. Since NADH:quinone oxidoreductase activity was not affected at these concentrations of EIPA, we conclude that H^+ and Na^+ transports were both decoupled from the oxidoreductase activity.

II.3.3.7 Complex I from *R. marinus* has two energy transducing sites

We have previously proposed a model for energy transduction by complex I from *R. marinus*, in which we suggested the presence of two energy transducing sites, one involving Na^+ independent H^+

translocation and the other processing Na^+ dependent H^+ translocation [7]. This model was based on the observations that complex I from *R. marinus* performs NADH:quinone oxidoreduction and H^+ and Na^+ transports and that Na^+ neither influences the catalytic activity nor is necessary for H^+ transport. However the presence of Na^+ increases H^+ transport. The studies here described not only corroborate the presence of two proton translocations sites, but also showed that the two sites have different inhibition profiles for EIPA. In the absence of EIPA, the two H^+ translocations sites (Na^+ dependent and independent) are perfectly operational; NADH-driven H^+ transport occurs, as well as an opposite NADH-driven Na^+ transport is performed. In the presence of 10 μM EIPA, the antiporter site is partially inhibited (50 %) whereas the sodium independent H^+ translocating site is not affected. These findings explain why, in these conditions, the $\Delta[Na^+]_{out}$ decreases in a higher amplitude when compared to ΔpH (30% for ΔpH_{out} and 50% for ΔpNa^+, Table II.3.1). In the presence of Na^+, 20 μM EIPA, causes 70-80% of inhibition of H^+ and Na^+ transports. However, in the absence of Na^+, the H^+ transport is only ~50% inhibited at the same concentration of EIPA, reflecting the inhibition profile of the Na^+ independent H^+ translocating site. The two different inhibition profiles for H^+ translocation, in a concentration range in which the oxidoreduction activity is not affected, reinforce the hypothesis that two H^+ translocating sites are present in complex I from *R. marinus.*

It was reported that in some situations, EIPA could also act as a protonophore, even stronger than CCCP [13-15]. This effect is

completely absent in our studies. EIPA did not influence NADH-driven formation of ΔΨ in the concentration range used for the translocation studies (0 to 20 µM). This conclusion was further supported by the ^{23}Na-NMR experiments in the presence of EIPA, in which an increase of Na^+ translocation was not observed, as expected for a protonophore [7]; on the contrary it worked as a Na^+ transport inhibitor.

II.3.3.8 Complex I transduces energy by an indirect coupling mechanism

Strong indications for the coupling mechanism for energy transduction in complex I have recently been provided by the structural data of the membrane domain of complex I from *E. coli* and of integral complex I from *T. thermophilus* and mitochondria of *Y. lipolitica*. The structures of the bacterial enzymes suggest that 3 protons may be translocated by an indirect coupling mechanism, involving conformational changes from near the quinone binding site to the antiporter subunits, with the participation of a long amphipatic α-helix, which spans over almost the entire membrane domain and may work as a transmission element [16]. In order to achieve the stoichiometry of 4 translocated H^+ per NADH molecule, proposed for complex I [17], an additional H^+ translocating site was considered. It would operate through a redox-driven mechanism, representing a direct coupling mechanism in which H^+ uptake would occur upon reduction of cluster N2 [18]. This hypothesis further suggested that the same process for H^+ uptake would be operational for H^+

translocation and for quinone reduction. Such a possibility seems unlikely in the case of complex I from *R. marinus* since the catalytic reaction and H^+ translocation, irrespective of the dependence of Na^+, can be decoupled, *i.e.* catalytic activity was not affected at certain conditions that highly inhibited H^+ translocation. The recent structural data [16, 19] do not exclude the possibility that the translocation of the fourth proton may be driven by conformational change. Moreover, based on our results and the available structures, we may speculate that EIPA could interfere with the transmission element hampering the required conformational changes, that follows the redox reaction, to propagate to the pumping site(s). In this case the energy released by the catalytic reaction would not be transduced into a transmembrane difference of electrochemical potential but instead it would be lost as heat. The possible decoupling of the activities of complex I could be physiologically relevant in a situation in which the cell would need to regenerate NADH, but would already reached its maxima of the transmembrane difference of electrochemical potential.

II.3.4 Conclusion

We propose that complex I from *R. marinus* has two types of energy transducing sites, both dependent on menaquinone reduction and both promoting H^+ translocation. However, one of the sites requires Na^+ for its operation. Complex I from *R. marinus* also promotes Na^+ transport in the opposite direction to that of H^+

transport. Furthermore, different EIPA inhibition profiles were observed for the two coupling sites. Taken together all data, we can hypothesize that one type of coupling site may work as a proton pump, while the other may operate by a Na^+/H^+ antiporter mechanism. In both cases energy transduction occurs through indirect coupling, most probably by conformational changes. These conclusions are the agreement with the obtained complex I structural data and deepen the clarification of the coupling mechanism for energy transduction by complex I.

II.3.5 Acknowledgments

We greatly appreciated the discussions and critical reading by M. Teixeira, A. S. Fernandes, A. Konstatinov and S. Todorovic. We thank J. Carita, ITQB for cell growth. We thank Rita Ventura and Filipa L. Sousa for the synthesis of DMN. We acknowledge CERMAX at ITQB and Rede Nacional de RMN for access to the facilities (REDE/1517/RMN/2005). EPR spectrophotometer was funded by FCT (FCT-REEQ/336/BIO/2005). A. P. Batista is recipient of a grant from Fundação para a Ciência e a Tecnologia (SFRH/BD/25288/2005). This project and B. C. Marreiros fellowship were funded by Fundação para a Ciência e a Tecnologia (PTDC/QUI-BIQ/100302/2008 to M.M.P.).

II.3.6 References

1. Pereira, M. M., Carita, J. N. & Teixeira, M. (1999) Membrane-bound electron transfer chain of the thermohalophilic bacterium *Rhodothermus marinus*: characterization of the iron-sulfur centers from the dehydrogenases and investigation of the high-potential iron-sulfur protein function by in vitro reconstitution of the respiratory chain, *Biochemistry.* 38, 1276-83.

2. Watters, C. (1978) A one-step biuret assay for protein in the presence of detergent, *Analytical biochemistry.* 88, 695-8.

3. Fernandes, A. S., Sousa, F. L., Teixeira, M. & Pereira, M. M. (2006) Electron paramagnetic resonance studies of the iron-sulfur centers from complex I of *Rhodothermus marinus*, *Biochemistry.* 45, 1002-8.

4. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid, *Analytical biochemistry.* 150, 76-85.

5. Apell, H. J. & Bersch, B. (1987) Oxonol VI as an optical indicator for membrane potentials in lipid vesicles, *Biochimica et biophysica acta.* 903, 480-94.

6. Briskin, D. P. & Reynolds-Niesman, I. (1991) Determination of H/ATP Stoichiometry for the Plasma Membrane H-ATPase from Red Beet (Beta *vulgaris L*.) Storage Tissue, *Plant Physiol.* 95, 242-250.

7. Batista, A. P., Fernandes, A. S., Louro, R. O., Steuber, J. & Pereira, M. M. (2010) Energy conservation by *Rhodothermus marinus* respiratory complex I, *Biochimica et biophysica acta*.

8. Delort, A. M., Gaudet, G. & Forano, E. (2002) 23Na NMR study of *Fibrobacter succinogenes* S85: comparison of three chemical shift reagents and calculation of sodium concentration using ionophores, *Analytical biochemistry.* 306, 171-80.

9. Stolpe, S. & Friedrich, T. (2004) The *Escherichia coli* NADH:ubiquinone oxidoreductase (complex I) is a primary proton pump but may be capable of secondary sodium antiport, *The Journal of biological chemistry.* 279, 18377-83.

10. Nakamaru-Ogiso, E., Seo, B. B., Yagi, T. & Matsuno-Yagi, A. (2003) Amiloride inhibition of the proton-translocating NADH-quinone oxidoreductase of mammals and bacteria, *FEBS letters.* 549, 43-6.

11. Damiano, E., Bassilana, M., Rigaud, J. L. & Leblanc, G. (1984) Use of the pH sensitive fluorescence probe pyranine to monitor internal pH changes in *Escherichia coli* membrane vesicles, *FEBS letters.* 166, 120-4.

12. Dlaskova, A., Hlavata, L., Jezek, J. & Jezek, P. (2008) Mitochondrial Complex I superoxide production is attenuated by uncoupling, *The international journal of biochemistry & cell biology.* 40, 2098-109.

13. Davies, K. & Solioz, M. (1992) Assessment of uncoupling by amiloride analogs, *Biochemistry.* 31, 8055-8.

14. Kleyman, T. R. & Cragoe, E. J., Jr. (1988) Amiloride and its analogs as tools in the study of ion transport, *The Journal of membrane biology.* 105, 1-21.

15. Dubinsky, W. P., Jr. & Frizzell, R. A. (1983) A novel effect of amiloride on H+-dependent Na+ transport, *The American journal of physiology.* 245, C157-9.

16. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I, *Nature.* 465, 441-5.

17. Galkin, A. S., Grivennikova, V. G. & Vinogradov, A. D. (1999) H+/2e-stoichiometry in NADH-quinone reductase reactions catalyzed by bovine heart submitochondrial particles, *FEBS letters.* 451, 157-61.

18. Berrisford, J. M. & Sazanov, L. A. (2009) Structural basis for the mechanism of respiratory complex I, *The Journal of biological chemistry.* 284, 29773-83.

19. Hunte, C., Zickermann, V. & Brandt, U. (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I, *Science.* 329, 448-51.

II.4 Complex I from *E. coli* devoid of NuoL subunit

This section is based on the following publication:

Marreiros BC, Batista AP, Pereira MM (2014) "Respiratory complex I from *Escherichia coli* does not transport Na^+ in the absence of its NuoL subunit" FEBS Letters, 588, 23, 4520-4525.

DOI:10.1016/j.febslet.2014.10.030

Author contribution:

B.C.M. performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

II.4 Complex I from *E. coli* devoid of NuoL subunit

II.4.1 Summary

We investigated H^+ and Na^+ transport by complex I from *E. coli* devoid of the NuoL subunit, which is probably part of the ion translocating machinery. We observed that complex I devoid of the NuoL subunit still translocates H^+, although to a smaller extension than the complete version of complex I, but does not transport Na^+. Our results unequivocally reinforce the observation that *E. coli* complex I transports Na^+ in the opposite direction to that of the H^+ and show that NuoL subunit is involved in the translocation of both ions by complex I.

II.4.2 Materials and Methods

II.4.2.1 PCR

We use an *E. coli* strain devoid of NuoL obtained from Keio collection [1]. The absence of *nuoL* gene from *E. coli* genome was verified by PCR. The selected primers were complementary to the DNA sequence located upstream and downstream of the *nuoL* gene. The amplified fragments were sequenced and their sizes were verified by electrophoresis in an agarose (1 %) gel. In this way the absence of the *nuoL* gene and the presence of the kanamycin cassette were confirmed.

II.4.2.2 Cell growth and membrane vesicles preparation

E. coli K12-MG1655 and *E. coli* containing complex I devoid of NuoL [1] were grown microaerophilicaly in LB medium at pH 7.0 and 37 °C. The cells were harvested in late exponential phase, to maximize the expression of complex I [2], suspended in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4, 25 mM Na_2SO_4 and disrupted in a French Pressure cell at 6,000 psi. The membrane vesicles were obtained by ultracentrifugation of the broken cells (200,000 *g*, 2 h, 4 °C) followed by re-suspension in the previous buffer. Protein concentration was determined using the Biuret method modified for membrane proteins [3].

II.4.2.2 Evaluation of complex I assembly

Assembly of complex I was evaluated by Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE) using a gradient gel (5-13%). BN-PAGE was carried out as described in [4]. *In-gel* Complex I activity was analyzed by NADH:Nitro blue tetrazolium (NBT) oxidoreductase activity. The gel was incubated in 50 mM potassium phosphate pH 7.0 containing 0.2 mg/mL NBT and 0.1 mg/mL NADH for 15-20 min, room temperature.

EPR spectra of solubilized membranes, reduced with 5 mM K_2-NADH in the presence of 5 mM KCN, were acquired on a Bruker EMX spectrometer equipped with an Oxford instruments ESR-900 continuous-flow helium cryostat, at 10 K, microwave frequency: 9.38 GHz and microwave power 2.4 mW.

Complex I activity, NADH:quinone oxidoreductase, was monitored at 340 nm (ε = 6220 $M^{-1}cm^{-1}$) and 25 ^{o}C, on a Shimadzu UV-1800 spectrophotometer inside of an anaerobic chamber. The reaction mixture contained membrane vesicles in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4 , 25 mM Na_2SO_4, 100 µM quinone, 2,3-dimethyl-1,4-naphthoquinone (DMN) or decylubiquinone (DUQ), and 100 µM K_2-NADH or deamino-NADH. DMN was synthesized as described in [5]. The obtained data are the average of at least three independent assays.

II.4.2.3 Characterization of the membrane vesicles

The orientation of membrane vesicles was investigated by the deamino-NADH:$K_3[Fe(CN)_6]$ oxidoreductase activity, in the absence or in the presence of n-Dodecyl-beta-Maltoside (DDM) detergent. The activity was monitored, at 420 nm (ε = 1020 $M^{-1}cm^{-1}$), 25 °C, on a Shimadzu UV-1800 spectrophotometer. The reaction mixture contained membrane vesicles or solubilized membrane vesicles in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4 , 25 mM Na_2SO_4, 250 µM $K_3[Fe(CN)_6]$ and 250 µM deamino-NADH. Solubilized membranes were obtained by stirring an aliquot of membrane vesicles with 2 g DDM/1 g membrane protein for 2 h, 4 °C. The internal volume of the membrane vesicles was determined by EPR spectroscopy, using TEMPO in the external medium and potassium chromium(III) oxalate as a quencher [6]. The NADH oxidase activity was monitored at 340 nm (ε = 6220 $M^{-1}cm^{-1}$) and 25 °C, on an OLIS upgraded Aminco DW2 dual wavelength spectrophotometer. The reaction mixture contained membrane

vesicles in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4 , 25 mM Na_2SO_4 and 100 µM K_2-NADH. Oxygen consumption was monitored at 25 °C on a Clark type oxygen electrode (Oxygraph from Hansatech instruments). The reaction was started by the addition of 4 mM K_2-NADH. When referred, KCN (2 mM) was added.

ΔΨ generation was monitored following oxonol VI absorption ($A_{628nm} - A_{587nm}$) at 25 °C, on an OLIS upgraded Aminco DW2 dual wavelength spectrophotometer [7]. The integrity of the vesicles was verified by NADH-driven ΔΨ formation and was performed by addition of 4 mM K_2-NADH to the vesicles. When referred, 100 µM CCCP, 2 mM KCN, 200 µM DMN or 200 µM DUQ were added.

II.4.2.4 Detection of ΔpH

ΔpH generation was investigated by fluorescence spectroscopy, following the quenching in ACMA fluorescence intensity ($\lambda_{excitation}$=410 nm, $\lambda_{emission}$=480 nm) at 25 °C, on a Varian Cary Eclipse spectrofluorimeter. The reaction mixture contained membrane vesicles in 2.5 mM HEPES-Tris pH 7.5, 5 mM K_2SO_4 , 25 mM Na_2SO_4 with 5 mM Mg_2SO_4, 100 µM DMN, 5 µM valinomycin, 2 µM ACMA and 400 µM K_2-NADH. When referred, 10 µM CCCP or 2 mM KCN were added prior to the addition of the substrate. The output was calculated with a moving average of 15 points.

II.4.2.5 Na^+ transport

Na^+ transport was monitored by ^{23}Na-NMR spectroscopy. NMR spectra were recorded on a Bruker Avance II 500 MHz spectrometer,

18 °C, operating at 132 MHz for ^{23}Na. Experiments were performed as described in [6]. Methylenephosphonate (4.5 mM) was used as a shift reagent for the Na^+ signal of the suspension medium. On average, 400 µL of membrane vesicles containing 27 mg of membrane protein were used in each NMR experiment in a 5 mm diameter tube in a total volume of 500 µL. A capillary tube containing the shift reagent dysprosium (III) tripolyphosphate (22 mM), was used in all experiments as external reference. Sodium concentration outside the membrane vesicles was determined by integrating the resonance frequency peak using the integration of the resonance frequency peak of Na^+ in the presence of $Dy(PPPi)_2^{7-}$ (inside the capillary) as reference. Spectra were obtained upon addition of 4 mM K_2-NADH to membrane vesicles or to membrane vesicles which were previously incubated with CCCP (10 µM), KCN (10 mM) or DMN (200 µM).

II.4.3 Results and Discussion

II.4.3.1 Functional assembly of complex I devoid of NuoL

The assembly of complex I devoid of NuoL was verified by BN-PAGE using solubilized membranes. For comparison, membranes from *E. coli* K12 containing the entire complex I were subjected to the same protocol. Bands compatible with the expected molecular masses for each complex I were observed in the two membrane preparations. The bands from both preparations presented NADH:NBT oxidoreductase activity, indicating the presence of a NADH

dehydrogenase (Figure II.4.1). The mass of the bands and their reactivity with NADH evidence the presence of complex I in the membranes of *E. coli* containing complex I devoid of NuoL.

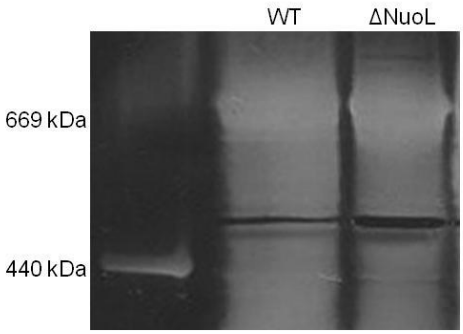


Figure II.4.1: WT and ΔNuoL Complex I is assembled and active. Blue Native Polyacrylamide Gel Electrophoresis, using a gradient gel (5-13%), of solubilized membranes of *E. coli* K12 (WT) and *E. coli* ΔNuoL strains and in gel activity (NADH:Nitrobluetetrazolium (NBT) oxidoreductase activity). WT and ΔNuoL presents bands compatible to the expected molecular mass of complex I (~540±50kDa and ~474±50kDa, respectively). NADH:NBToxidoreductase activity was observed for both strains.

The membrane vesicles from *E. coli* containing complex I devoid of NuoL showed NADH:DMN, deamino-NADH:DMN, NADH:DUQ and deamino-NADH:DUQ oxidoreductase activities of 254.5±15.5 nmol $NADH.mg_{protein}^{-1}.min^{-1}$, 38.7±3.7 nmol $dNADH.mg_{protein}^{-1}.min^{-1}$, 395.2±4.5 nmol $NADH.mg_{protein}^{-1}.min^{-1}$ and 47.1±1.2 nmol $dNADH.mg_{protein}^{-1}.min^{-1}$, which compare, under the same conditions, to 219.8±21.9 nmol $NADH.mg_{protein}^{-1}.min^{-1}$ 52.5±8.7 nmol $dNADH.mg_{protein}^{-1}.min^{-1}$, 428.5±13.6 nmol $NADH.mg_{protein}^{-1}.min^{-1}$, 119.7±18.0 nmol $dNADH.mg_{protein}^{-1}.min^{-1}$ for membrane vesicles containing the entire complex I. The assembly of complex I devoid of NuoL was further supported by the observation that solubilized membrane vesicles containing this complex I presented EPR signals

upon reduction with NADH in the presence of KCN compatible with the presence of the N2 iron-sulfur cluster (Figure II.4.2).

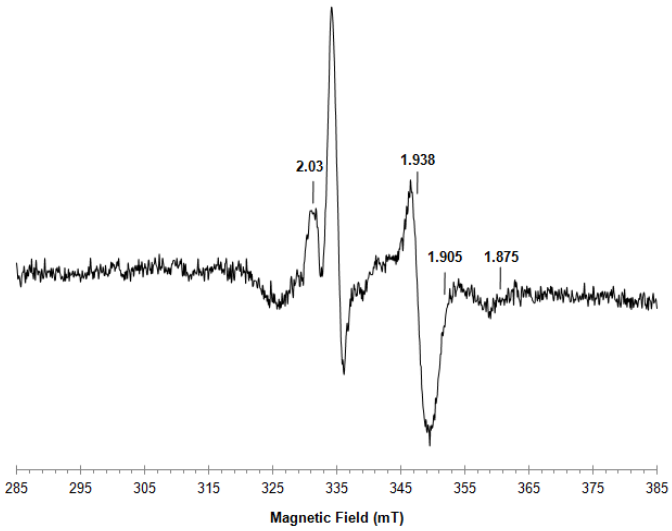


Figure II.4.2: EPR spectrum of reduced solubilized membranes from *E. coli* expressing Complex I devoid of NuoL subunit I. The solubilized membranes were reduced with 5 mM K_2-NADH in the presence of 5 mM KCN. The spectrum was obtained at 10 K, microwave frequency, 9.38 GHz, microwave power 2.4 mW and modulation amplitude of 1 mT.

II.4.3.2 Characterization of the membrane vesicles from *E. coli* containing complex I devoid of NuoL

Membrane vesicles containing complex I devoid of NuoL and those containing the entire complex I showed deamino-NADH:$K_3[Fe(CN)_6]$ oxidoreductase activities of 438.8±25.4 nmol dNADH.$mg_{protein}^{-1}$.min^{-1} and 418.2±32.2 nmol dNADH.$mg_{protein}^{-1}$.min^{-1}, respectively. The former presented an inside-out orientation of 85% and the latter of 82%. Both membrane vesicles had an internal volume of 1 µL.mg^{-1}, in agreement with previous results [8]. NADH oxidation by the respiratory chain was 165.0±16.6 nmol NADH.$mg_{protein}^{-1}$.min^{-1}

and 160.86±22.9 nmol NADH.$mg_{protein}^{-1}$.min^{-1} for the membrane vesicles containing complex I devoid of NuoL and those containing the entire complex I. The activity was 63% and ~55% inhibited, respectively, in the presence of 0.5 uM piericidin A, the typical inhibitor of *E. coli* complex I. Both preparations contained tight vesicles allowing a stable formation of ΔΨ, when K_2-NADH was added, in the presence of oxygen. The establishment of ΔΨ was abolished in the presence of the protonophore CCCP or inhibited in the presence of KCN (Figure II.4.3, panel A and B). Establishment of ΔΨ was also observed in the presence of KCN, for both preparations, upon addition of K_2-NADH when DMN or DUQ were present (Figure II.4.3, panel C and D). This observation indicates the establishment of ΔΨ by the activity of complex I.

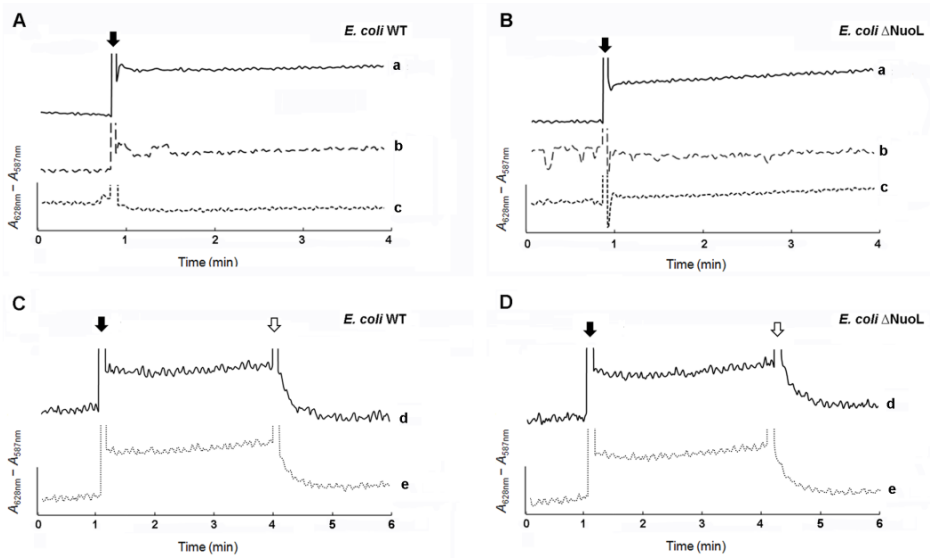


Figure II.4.3: NADH-driven ΔΨ generation by *E. coli* membrane vesicles. NADH-driven ΔΨ generation by *E. coli* membrane vesicles. ΔΨ generation was detected

following oxonol VI absorption ($A_{628} - A_{587}$) at 25 °C using membrane vesicles of *E. coli*containing the entire complex I (Panel A and C) and *E. coli* containing complex I devoid of NuoL (Panel B and D) upon addition of 400 µM K_2-NADH (indicated by a black arrow) and 100 µM CCCP (indicated by an white arrow). (a) NADH-driven $\Delta\Psi$ generation of respiratory chain; (b) NADH-driven $\Delta\Psi$ generation of respiratory chain inhibited by 2 mM KCN; (c) NADH-driven $\Delta\Psi$ generation of respiratory chain in the presence of 100 µM CCCP; (d) NADH-driven $\Delta\Psi$ generation of respiratory chain in the presence of 2 mM KCN and 200 µM DMN; (e) NADH-driven $\Delta\Psi$ generation of respiratory chain in the presence of 2 mM KCN and 200 µM DUQ. The represented data are the average of at least three independent assays.

II.4.3.3 H^+ and Na^+ transport by Complex I from *E. coli* containing complex I devoid of NuoL

H^+ transport by complex I was monitored by following the quenching in 9-amino, 6-chloro, 2-methoxyacridine (ACMA) fluorescence intensity, using the two preparations of membrane vesicles, one containing the entire complex I and another containing complex I devoid of NuoL. H^+ transport by complex I, upon addition of K_2-NADH was observed for both preparations (Figure II.4.4). H^+ transport was not observed in the presence of CCCP. Considering an equivalent relative amount of complex I in the two preparations based on their deamino-NADH:DMN oxidoreductase activity, a smaller H^+ translocation by complex I devoid of NuoL can be estimated. We observed that the ACMA quenching in the presence of the membranes containing complex I devoid of NuoL upon addition of NADH is less than the ACMA quenching in the presence of membranes with entire complex I, under the same conditions. This result is in agreement with other reported works, in which the NuoL subunit was deleted, C-terminal truncated or point mutated. In all those cases, complex I presented less H^+ translocation activity [9-13].

We have shown previously that complex I from *E. coli* has Na^+/H^+ antiporter activity [8]. Thus Na^+ transport by complex I devoid

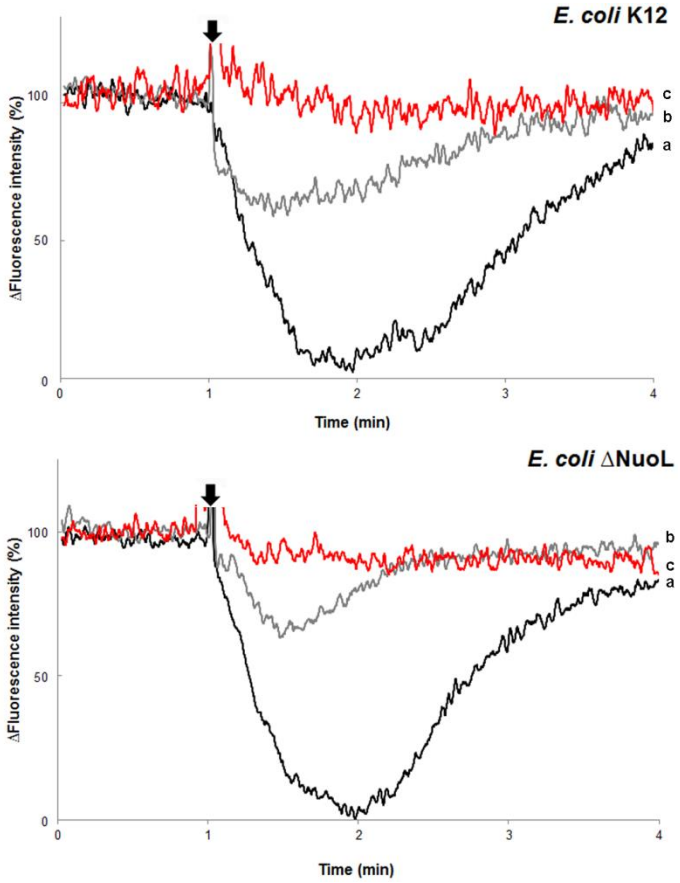


Figure II.4.4: ΔpH generation by *E. coli* membrane vesicles. Quenching of ACMA fluorescence intensity ($\lambda_{excitation}$ = 410 nm, $\lambda_{emission}$ = 480 nm) at 25 °C using membrane vesicles of *E. coli* containing the entire complex I (Panel A) and *E. coli* containing complex I devoid of NuoL (Panel B) upon addition of 400 µM K_2-NADH (indicated by an arrow). The reaction mixture contained membrane vesicles in 2.5 mM HEPES–Tris pH 7.5, 25 mM Na_2SO_4, 5 mM $K_2SO_4$with 5 mM Mg_2SO_4, 150 µM DMN, 5 µM valinomycin and 2 µM ACMA. (a) No further addition; (b) r in the presence of 2 mM KCN and 200 µM DMN (complex I); (c) in the presence of 100 µM CCCP. The represented data are the average of at least three independent assays.

of NuoL was investigated using ^{23}Na-NMR spectroscopy (Figure II.4.5). For comparison and as a control, Na^+ transport by inverted membrane

vesicles containing the entire complex I was monitored under the same conditions. Briefly, upon addition of NADH Na^+ transport to the outside of these vesicles, in the opposite direction of H^+ transport, was observed. Na^+ transport was not observed in the presence of KCN, the inhibitor of heme-copper oxygen reductases, the terminal reductases of the electron transport chain. Na^+ transport was recovered in these conditions by the presence of DMN, *i.e* by NADH:DMN oxidoreductase activity. In the presence of CCCP, Na^+ transport by NADH:quinone oxidoreductase activity increased, indicating that the transport is a primary event. These results were equal to those obtained before [6, 8]. In the case of *E. coli* membrane vesicles containing complex I devoid of the NuoL, under all experienced conditions, no Na^+ transport was detected (Figure II.4.5).

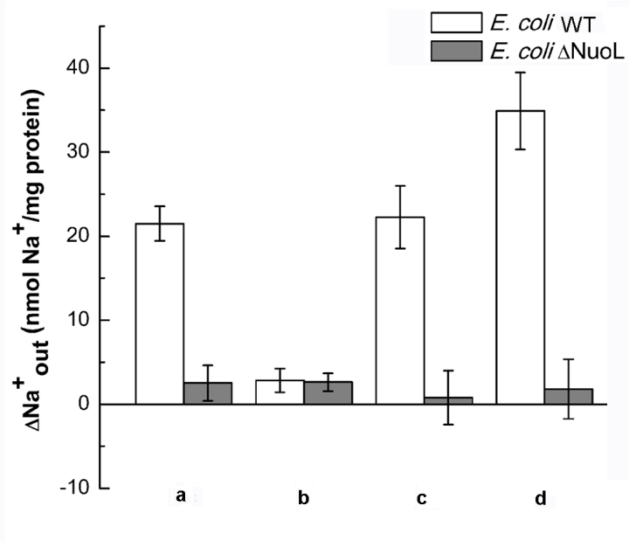


Figure II.4.5: Na^+ transport by *E. coli* membrane vesicles. $^{23}Na^+$ transport by *E. coli* membrane vesicles. ^{23}Na-NMR spectroscopy was used to monitor Na^+ transport, 5 min after K_2-NADH addition, using membrane vesicles of *E. coli*containing the entire complex I (white columns) and *E. coli* containing complex I devoid of NuoL (gray columns). (a) No further addition; (b) in the presence of 10 mM KCN; (c) in the

presence of 10 mM KCN and 200 µM DMN; (d) in the presence of 10 mM KCN, 200 µM DMN and 10 µM CCCP. The represented data are the average of at least three independent assays.

II.4.3.4 The role of NuoL subunit

In this work we investigated H^+ and Na^+ transports by *E. coli* complex I devoid of NuoL. We aimed at (i) reinforcing our previous observations that complex I from *E. coli* translocates Na^+ ions in the opposite direction of the H^+, *i.e*. that it performs Na^+/H^+ antiporter activity and (ii) to investigate the role of NuoL in ion transport.

NuoL subunit is located at the distal end of the membrane arm in relation to the base of the peripheral arm (Figure II.4.6). It is one of the antiporter-like subunits suggested to contain one ion translocating site. In addition, it contains a unique C-terminal facing the cytoplasmic side, containing a ~110 Å HL spanning laterally to itself, NuoM and N subunits, which may have a determinant role in energy transduction. (Figure II.4.6). The C-terminal ends in a TM helix (TM16) anchored at the interface of NuoN (TM6 and TM7 helices), NuoK (TM1 helix and C-terminal) and NuoJ (TM2 helix) subunits [14]. Since this anchoring helix is on the side of the membrane domain its removal is not expected to destabilize significantly the domain and thus the complex. Complex I devoid of NuoL contains three of the proposed ion translocating sites, NuoM, N and the consortium composed of NuoH, J and K subunits [15].

Four independent studies using complex I devoid of NuoL were carried out before, being apparently contradictory [9, 10, 12, 16]. Two

of these studies suggested that complex I devoid of NuoL does not assemble or presents a low expression and/or activity [10, 16]. On the

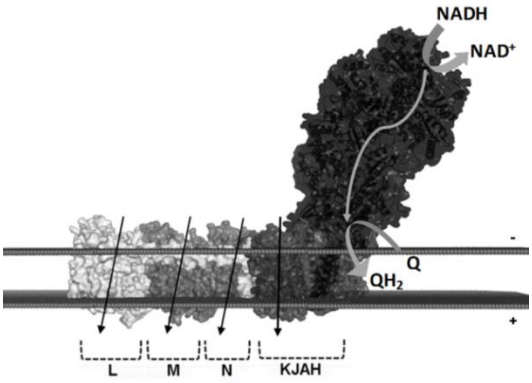


Figure II.4.6: Schematic representation of respiratory complex I and its ion translocation sites (Adapted from PDB:3M9S). Bacterial complex I is in general constituted by 14 subunits designated NuoA to N. The subunits are arranged in an L-shaped structure, consisting of a peripheral (NuoB-G and I) and a membrane (NuoA, H and J-N) arms. The peripheral part contains the catalytic sites where NADH is oxidized and quinone reduced. The membrane domain includes the ion translocating machinery composed of seven subunits, almost linearly arranged as NuoH, A, J, K, N, M and L. Complex I contains four proposed ion translocating sites, NuoL M, N (light gray and white subunits) and the consortium composed of NuoH, J and K subunits (dark gray subunits) [15]. The black arrows indicate the approximate sites

other hand the two other works showed that complex I without NuoL can assemble and is active [9, 12]. The experiments performed were differently designed, including the use of different strains, which may explain first sight discrepancies regarding the assembly of complex I devoid of NuoL. Belevich *et al* used *E. coli* GR70N and evaluated the production/assemble of complex I by the presence of deamino-NADH: hexaammineruthenium (III) chloride (HAR) oxidoreductase activity [16]. The authors recognized that the used activity assay may not be always an accurate parameter. Nevertheless, the lack of deamino-NADH:HAR oxidoreductase activity associated to the impaired growth

of *E. coli* GRL3 in minimal medium led the authors to conclude that complex I was not assembled if gene *nuoL* was deleted from the chromosome of *E. coli* GRL3 [16] . Torres-Bacete and coworkers used a NuoL knockout mutant from *E. coli* MC4100 [17]. They did not observe deamino-NADH oxidase or deamino-NADH:Q oxidoreductase activities. Still, other subunits of complex I, including NuoM subunit, were present in the membranes as observed by immunobloting assays. The authors suggested that the absence of activities by that mutant could be explained by the incorrect assembly of complex I [17]. Steimle *et al* worked with *E. coli* BW25113 and were able to purify complex I devoid of NuoL subunit [12]. Both the entire complex I and the one devoid of NuoL presented similar EPR spectra, when reduced with NADH. The complex devoid of NuoL was shown to be active and to translocate H^+, although to a smaller extension [12].

We observed that complex I devoid of NuoL subunit from *E. coli* assembles and is active, presenting lower H^+ translocation, when compared to the entire complex. In addition no Na^+ transport was observed. It can thus be hypothesized that i) NuoL subunit is the site in which Na^+ and part of H^+ transports occur, ii) NuoL influences ions transport by other subunits or iii) both.

Subunits NuoL, M and N are homologous to subunits MrpA and MrpD of the so-called Mrp Na^+/H^+ antiporters. Recently, it was observed that cells from *B. subtilis* in which the genes coding for MrpA or MrpD have been deleted, could closely recover the respective growth profiles when expressing NuoL, M or N subunits [18, 19]. Previously, membrane vesicles of *E. coli* lacking Na^+/H^+ antiporters

were shown to exhibit a higher Na^+ uptake, when a C-terminal truncated version of NuoL subunit is expressed. This C-terminal truncated version of NuoL does not contain the HL or the second inverted repeat. Still, Na^+ uptake was also observed when this construct was reconstituted in liposomes [20]. All together these results reinforce that NuoL is involved in Na^+ transport, being possibly a Na^+/H^+ antiporter. Na^+ transport may not be exclusive to NuoL, since expression of NuoM or N also recovered the growth profiles of *B. subtilis* devoid of *mrpA* or *mrpD* genes. Thus NuoL may also influence the transport of this ion by other antiporter-like subunit(s) in the entire complex I.

In this work, complex I devoid of the NuoL subunit was observed to be able to transport H^+, although to a lower extent. These results are in agreement with the observed lower H^+/e stoichiometry in *E. coli* complex I devoid of the NuoL subunit [12] and in *Y. lipolytica* complex I devoid of the NuoL and M subunits [9]. In this case, the observation that the same stoichiometry was obtained irrespectively of the presence or absence of NuoM subunit may suggest that this subunit is not operating in complex I devoid of NuoL. Such interpretation supports the hypothesis that NuoL is also influencing the ion transport by other subunits. In this respect, a cooperative electrostatic coupling between titratable residues at the interface of the antiporter-like subunits has been recently suggested by theoretical calculations [21].

The role of the extended C-terminal of the NuoL subunit has been largely discussed from two different perspectives: one suggests

its function as a coupling element and the other as a stabilizing component, keeping the three antiporter-like subunits together [9, 12, 13, 16, 17, 22]. We have predicted the presence of an antiporter-like subunit with an extended C-terminal in several complexes related to complex I, independently of the number of the antiporter-like subunits present [23, 24] (see sections II.1 and II.2). These complexes include the energy-converting hydrogenases and formate hydrogen lyases-1, which contain only one antiporter-like subunit. Such prediction questions an exclusive stabilizing role for the HL, since only one antiporter-like subunit is present [23].

In conclusion, our transport analyses showed that complex I devoid of NuoL is unable to transport Na^+, reinforcing our previous observation that *E. coli* complex I transports Na^+ ions in the opposite direction to the transport of H^+, the coupling ion. Our H^+ translocation assays indicate that complex I devoid of NuoL is less efficient than the entire complex I, in agreement with previous observations [12]. All together, these observations suggest that NuoL can work as a Na^+/H^+ antiporter, and additionally it may also influence the ion transport activity of other subunits.

II.4.4 Acknowledgments

We thank Cristina Paulino for technical discussions on fluorescence spectroscopy and João Carita for cell growth. The NMR spectrometers are part of The National NMR Facility, supported by Fundação para a Ciência e a Tecnologia (RECI/BBB-BQB/0230/2012).

A.P.B. is recipient of a grant from Fundação para a Ciência e a Tecnologia (SFRH/BPD/80741/2011). The project was funded by Fundação para a Ciência e a Tecnologia (PTDC/BBB-BQB/2294/2012 to M.M.P.). The work was supported by Fundação para a Ciência e a Tecnologia through grant # PEst-OE/EQB/LA0004/2011.

II.4.5 References

1. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. & Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, *Molecular systems biology.* 2, 2006 0008.

2. Unden, G. & Bongaerts, J. (1997) Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors, *Biochimica et biophysica acta.* 1320, 217-34.

3. Watters, C. (1978) A one-step biuret assay for protein in the presence of detergent, *Analytical biochemistry.* 88, 695-8.

4. Wittig, I., Braun, H. P. & Schagger, H. (2006) Blue native PAGE, *Nature protocols.* 1, 418-28.

5. Kruber, O. (1929) Über das 2.3-Dimethyl-naphthalin im Steinkohlenteer, *European Journal of Inorganic Chemistry.* 62, 3044–3047.

6. Batista, A. P., Fernandes, A. S., Louro, R. O., Steuber, J. & Pereira, M. M. (2010) Energy conservation by *Rhodothermus marinus* respiratory complex I, *Biochimica et biophysica acta*.

7. Apell, H. J. & Bersch, B. (1987) Oxonol VI as an optical indicator for membrane potentials in lipid vesicles, *Biochimica et biophysica acta.* 903, 480-94.

8. Batista, A. P. & Pereira, M. M. (2011) Sodium influence on energy transduction by complexes I from *Escherichia coli* and *Paracoccus denitrificans*, *Biochimica et biophysica acta.* 1807, 286-92.

9. Drose, S., Krack, S., Sokolova, L., Zwicker, K., Barth, H. D., Morgner, N., Heide, H., Steger, M., Nubel, E., Zickermann, V., Kerscher, S., Brutschy, B., Radermacher, M. & Brandt, U. (2011) Functional dissection of the proton pumping modules of mitochondrial complex I, *PLoS biology.* 9, e1001128.

10. Nakamaru-Ogiso, E., Kao, M. C., Chen, H., Sinha, S. C., Yagi, T. & Ohnishi, T. (2010) The membrane subunit NuoL(ND5) is involved in the indirect proton pumping mechanism of *Escherichia coli* complex I, *The Journal of biological chemistry.* 285, 39070-8.

11. Erhardt, H., Steimle, S., Muders, V., Pohl, T., Walter, J. & Friedrich, T. (2012) Disruption of individual nuo-genes leads to the formation of partially assembled NADH:ubiquinone oxidoreductase (complex I) in *Escherichia coli*, *Biochimica et biophysica acta.* 1817, 863-71.

12. Steimle, S., Bajzath, C., Dorner, K., Schulte, M., Bothe, V. & Friedrich, T. (2011) Role of subunit NuoL for proton translocation by respiratory complex I, *Biochemistry.* 50, 3386-93.

13. Steimle, S., Willistein, M., Hegger, P., Janoschke, M., Erhardt, H. & Friedrich, T. (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I, *FEBS letters.* 586, 699-704.

14. Efremov, R. G., Baradaran, R. & Sazanov, L. A. (2010) The architecture of respiratory complex I, *Nature.* 465, 441-5.

15. Baradaran, R., Berrisford, J. M., Minhas, G. S. & Sazanov, L. A. (2013) Crystal structure of the entire respiratory complex I, *Nature.* 494, 443-8.

16. Belevich, G., Knuuti, J., Verkhovsky, M. I., Wikstrom, M. & Verkhovskaya, M. (2011) Probing the mechanistic role of the long alpha-helix in subunit L of respiratory Complex I from *Escherichia coli* by site-directed mutagenesis, *Molecular microbiology.* 82, 1086-95.

17. Torres-Bacete, J., Sinha, P. K., Matsuno-Yagi, A. & Yagi, T. (2011) Structural contribution of C-terminal segments of NuoL (ND5) and NuoM (ND4) subunits of complex I from *Escherichia coli*, *The Journal of biological chemistry.* 286, 34007-14.

18. Moparthi, V. K., Kumar, B., Mathiesen, C. & Hagerhall, C. (2011) Homologous protein subunits from *Escherichia coli* NADH:quinone oxidoreductase can

functionally replace MrpA and MrpD in *Bacillus subtilis*, *Biochimica et biophysica acta.* 1807, 427-36.

19. Moparthi, V. K., Kumar, B., Al-Eryani, Y., Sperling, E., Gorecki, K., Drakenberg, T. & Hagerhall, C. (2014) Functional role of the MrpA- and MrpD-homologous protein subunits in enzyme complexes evolutionary related to respiratory chain complex I, *Biochimica et biophysica acta.* 1837, 178-85.

20. Steuber, J. (2003) The C-terminally truncated NuoL subunit (ND5 homologue) of the Na+-dependent complex I from *Escherichia coli* transports Na+, *The Journal of biological chemistry.* 278, 26817-22.

21. Kaila, V. R., Wikstrom, M. & Hummer, G. (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I, *Proceedings of the National Academy of Sciences of the United States of America.* 111, 6988-93.

22. Galkin, A., Drose, S. & Brandt, U. (2006) The proton pumping stoichiometry of purified mitochondrial complex I reconstituted into proteoliposomes, *Biochimica et biophysica acta.* 1757, 1575-81.

23. Batista, A. P., Marreiros, B. C. & Pereira, M. M. (2013) The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes, *Biological chemistry.* 394, 659-66.

24. Marreiros, B. C., Batista, A. P., Duarte, A. M. & Pereira, M. M. (2012) A missing link between complex I and group 4 membrane-bound [NiFe] hydrogenases, *Biochimica et biophysica acta.* 1827(2), 198-209.

Chapter III

NDH-2 family

Chapter III – NDH-2 family

III.1 - NDH-2 family: Phylogenetic distribution, Structural diversity and Evolutionary divergences

This section is based on the following publication:

Marreiros BC, Sena FV, Sousa FM, Batista AP, Pereira MM (2016) "Type II NADH:quinone oxidoreductase family: Phylogenetic distribution, Structural diversity and Evolutionary divergences" Environmental Microbiology.

DOI: 10.1111/1462-2920.13351

Author contribution:

B.C.M. participated in the design of the study, performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

III.1 - NDH-2 family: Phylogenetic distribution, Structural diversity and Evolutionary divergences

III.1.1 Summary

Type II NADH:quinone oxidoreductases (NDH-2s) are membrane proteins, crucial for the catabolic metabolism, because they contribute to the maintenance of the NADH/NAD^+ balance. In several bacteria and protists, namely pathogenic, NDH-2s are the only enzymes performing respiratory NADH:quinone oxidoreductase activity. For this reason and for being considered absent in mammals, NDH-2s were proposed as suitable targets for novel antimicrobial therapies.

We correlate the available information on all characterized NDH-2s with sequence analyses. We selected all sequences of genes encoding NDH-2s from fully sequenced genomes present in the KEGG database. These genes were present in 61 % of the 1805 species belonging to Eukarya (83 %), Bacteria (60 %) and Archaea (32 %). Notably sequences from mammal species including humans were retrieved in our selection as NDH-2s.

The data obtained and the already available information allowed systematizing several properties of NDH-2s: (i) the existence of additional sequence motifs with putative regulatory functions, (ii) specificity towards NADH or NADPH and (iii) the type of quinone binding motif.

We observed that NDH-2 family distribution is not congruent with the taxonomic tree, suggesting different origins for the eukaryotic

sequences and possible lateral gene transfer among prokaryotes. We note the absence of genes coding for NDH-2 in anaerobic phyla and the presence of multiple copies in several genomes, specifically in cyanobacteria. These observations inspired us to propose a metabolic hypothesis for the appearance of NDH-2s.

III.1.2 Material and Methods

III.1.2.1 Database search and selection of NDH-2 family

The selection of the sequences composing the NDH-2 family was not trivial, because the members of the different families included in the two Dinucleotide Binding Domain Flavoproteins (tDBDF) superfamily are difficult to distinguish and many have been previously misannotated. Thus a careful selection of sequences had to be performed. Genes coding for proteins homologous to the members of the tDBDF superfamily were searched among the Eukarya, Bacteria and Archaea domains using protein BLAST (BLASTp) analysis tool running at KEGG´s (Kyoto Encyclopedia of Genes and Genomes) database platform [1, 2] (Figure III.1.1, step 1). This database was chosen since it contains data on only full sequenced genomes, which is a requirement to search for the presence or absence of genes in specific species. The information used was that available by September 2014 (this database was updated in section IV.1). Based on characterized members of the tDBDF superfamily [3] at least one protein sequence representative of each family was used as a query

for pBLAST (default parameters were used) (Table III.1.1). All amino acid sequences with E-value < 0.01 were accepted for analysis obtaining an initial data set of ~32,000 amino acid sequences (Figure III.1.1, step 2). Amino acid sequences and the respective information were retrieved using a homemade script on Microsoft Office Excel using Visual Basic for Applications (VBA) programming language. We filtered the data set by amino acid sequence length to exclude proteins that could be fused to other proteins. We selected the amino acid sequences comprising between 250 and 1000 or 1150 amino acid residues for Prokarya or Eukarya, respectively. We considered longer sequences in the case of Eukarya sequences because we took into account that some of these have a signal peptide localization at the N-terminal (mitochondrial or chloroplast signal peptides) or an amino acid sequence insertion (e.g. EF-hand motif).

We obtained a data set of ~29,500 amino acid sequences (Figure III.1.1, step 3). Then, we clustered the amino acid sequences according to their identity using the CD- HIT tool [4]. As the main goal here was to reduce the size of the dataset for further amino acid sequence alignment, we selected 50 % identity to cluster the ~29,500 amino acid sequences ensuring in this way that amino acid sequences from different families do not cluster together. We obtained a new data set of 5244 representative sequences and the respective clusters (Figure III.1.1, step 4). Afterwards, we aligned the data set using ClustalX 2.1 [5] and constructed the correspondent Neighbor-joining (NJ) dendrogram (see below, Figure III.1.3). In this dendrogram we were able to identify several branches belonging to different families

of the tDBDF superfamily, namely that designated before as NADH dehydrogenase (see below, Figure III.1.4, sequences highlighted in purple) in which the NDH-2 and sulfide dehydrogenase families (SQR and FCSD) are

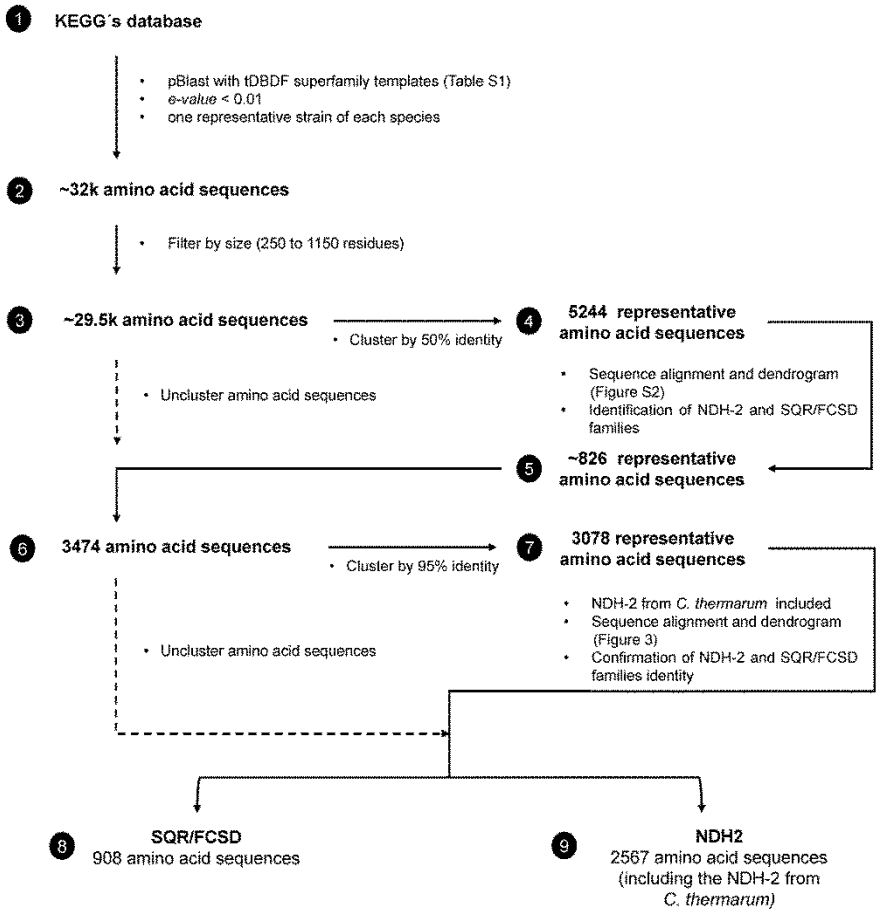


Figure III.1.1. Graphical representation of the procedure used for the selection of NDHs-2 dataset. SQR (sulfide:quinone oxidoreductase) and FCSD (sulfide:flavocytochrome *c* oxidoreductase also called flavocytochrome *c* sulfide dehydrogenase) were also selected with this procedure.

Table III.1.1. Proteins belonging to the tDBDF superfamily, whose sequences were used as queries for pBLAST search.

Group	Species	KEGG ID	UniProt ID	Locus tag	PDB
Flavin containing monooxygenase (MOX)	*Pseudomonas fluorescens*	pfo:Pfl01_2048	Q3KEL5	Pfl01_2048	-
	Aspergillus fumigatus	afm:AFUA_2G07680	E9QYP0	AFUA_2G07680	4B63, 4NZH, 4B64, *et al*
	Rhodococcus erythropolis	rer:RER_pREL1-02970	Q3L973	RER_pREL1-02970	-
2,4-dienoyl-CoA reductase (DCR)	*Escherichia coli K-12*	eco:b3081	P42593	b3081	1PS9
Adrenodoxin reductase (ADR)	*Bos taurus*	bta:282604	P08165	-	1CJC, 1E1M, 1E1K, *et al*
Glutamate synthase (GMS)	*Sus scrofa (pig)*	ssc:397109	Q28943	-	1GTE, 1H7X, 1GTH, *et al*
Alkylhydroperoxide reductase (AHR)	*Escherichia coli K-12*	eco:b0606	P35340	b0606	-
	Escherichia coli K-12	eco:b0888	P35340	b0888	1FL2, 4O5Q, 4YKG, *et al*
Disulfide reductase (DSR)	*Homo sapiens*	hsa:2936	P00390	-	3DK9, 3DK8, 3DJJ, *et al*
	Escherichia coli K-12	eco:b0116	P0A9P0	b0116	4JQ9, 4JDR
	Saccharomyces cerevisiae	sce:YFL018C	P09624	YFL018C	1V59, 1JEH
NADH ferredoxin reductase (NFR)	*Plasmodium falciparum*	pfa:PFI1170c	P61076	PFI1170c	4J56, 4J57, 4B1B
	Homo sapiens	hsa:9131	O95831	-	4BV6, 1M6I, 4LII, *et al*
	Pseudomonas monteilii	pmon:X969_10840	UPI00000B7951	X969_10840	-
	Pyrococcus furiosus	pfu:pf1197	Q8U1K9	PF1197	1XHC
NADH peroxidase/oxidase and disulfide CoA reductase (POR)	*Enterococcus faecalis*	efa:EF1932	Q833L5	EF1932	3OC4
	Streptococcus pyogenes	spy:spy_1150	Q99ZN6	SPy_1150	2BC0, 2BCP, 2BC1
	Staphylococcus aureus	sau:SA0831	UPI00000CAA73	SA0831	4EQS, 4EQW, 1YQZ, *et al*
NADH dehydrogenase (NDH)	*Saccharomyces cerevisiae*	sce:YML120C	P32340	YML120C	4G6H, 4G9K
	Staphylococcus aureus	sao:SAOUHSC_00878	Q2FZV7	SAOUHSC_00878	4XDB
	Caldalkalibacillus thermarum	-	F5L3B8	-	4NWZ
	Allochromatium vinosum	alv:Alvin_1092	Q06530	Alvin_1092	1FCD
	Acidithiobacillus ferrooxidans	afe:Lferr_1469	B5ES29	Lferr_1469	3SX6, 3SY4, 3SZF, *et al*

included [6]. We identified in this way 826 amino acid sequences (< 50 % identity) representatives of a data set of 3474 amino acid sequences prior to CD-HIT clustering (Figure III.1.1, step 6). To discriminate between NDH-2 and sulfide dehydrogenase families we performed the alignment of those 3474 sequences and corresponding NJ dendrogram (see below, Figure III.1.4). To ensure a robust identification of the amino acid sequences belonging to the two families (NDH-2 and sulfide dehydrogenase), we verified (a) the distribution of the 26 biochemically characterized NDH-2 (see below, Table III.1.2), (b) the information present in the KEGG´s database for randomly selected amino acid sequences, (c) the existence of associated PDB structures, including the three Sulfide:quinone oxidoreductase (SQR) [6-8], where the protein isolated from the archaeon *Acidianus ambivalens* was included since was shown to be a SQR instead of a NDH-2 [8] and the one sulfide:flavocytochrome *c* oxidoreductase (also called flavocytochrome *c* sulfide dehydrogenase, FCSD) [9] and (d) the presence of disulfide redox-active sites, a key feature of the sulfide dehydrogenase family. In this way we were able to clearly distinguish the two families. At the end we selected 2567 sequences as NDH-2, which we used to perform a taxonomic profiling and phylogenetic analyses (Figure III.1.1, step 9).

III.1.2.2 Sequence analysis

Multiple Sequence Alignments (MSA) were performed using ClustalX 2.1 (Protein weight matrix Gonnet, with Gap Opening 10 and Gap Extension 0.2) [5] and Mafft 7.187 (FFT-NS-2) [10]. The MSA were

Table III.1.2. List of biochemically characterized NDHs-2.

Fig. 3	Species	KEGG ID	UniProt ID	Locus tag	Size (aa)	PDB	References
B - 1a	*Saccharomyces cerevisiae*	sce:YNR074C	P52923	YNR074C	378	-	Wissing *et al*, 2004
B - 2	*Homo sapiens*	hsa:84883	Q9BRQ8	-	373	-	Marshall *et al*, 2005; Elguindy *et al*, 2015
B - 3a	*Neurospora crassa*	ncr:NCU09447	Q7S1W8	NCU09447	485	-	Carneiro *et al*, 2007
D1 - 4a	*Staphylococcus aureus*[1]	sau:SA0799	W8TN11	SA0799	354	-	Schurig-Briccio *et al*, 2014
D1 - 4b	*Staphylococcus aureus*[1]	sau:SA0802	Q2FZV7	SA0802	402	4XDB	Schurig-Briccio *et al*, 2014; Sena *et al*, 2015
D1 - 5	*Caldalkalibacillus thermarum*	-	F5L3B8	-	399	4NWZ	Heikal *et al*, 2014
D1 - 6	*Bacillus pseudofirmus*	bpf:BpOF4_04810	A7LKG4	BpOF4_04810	405	-	Liu *et al*, 2008
D2 - 7	*Methylococcus capsulatus*	mca:MCA1918	Q606U6	MCA1918	440	-	Cook *et al*, 2002
D2 - 8	*Escherichia coli*	eco:b1109	P00393	b1109	434	-	Bjorklof *et al*, 2000
D2 - 9	*Gluconobacter oxydans*	gox:GOX1675	Q5FQD1	GOX1675	409	-	Mogi *et al*, 2009
D4 - 10	*Mycobacterium tuberculosis*	mtu:Rv1854c	P95160	Rv1854c	463	-	Yano *et al*, 2006; 2014
D4 - 11	*Mycobacterium smegmatis*	msm:MSMEG_3621	O52473	MSMEG_3621	457	-	Vilcheze *et al*, 2005
D4 - 12	*Corynebacterium glutamicum*	cgl:NCgl1409	Q79VG1	NCgl1409	467	-	Nantapong *et al*, 2005
D4 - 13	*Agrobacterium tumefaciens*[2]	atf:Ach5_19080	UPI000618764E	Ach5_19080	421	-	Bernard *et al*, 2006
D5 - 1b	*Saccharomyces cerevisiae*	sce:YML120C	P32340	YML120C	513	4G6H; 4G9K	Velazquez and Pardo, 2001; Yamashita *et al*, 2007; Iwata et al, 2012; Feng et al, 2012
D5 - 1c	*Saccharomyces cerevisiae*	sce:YDL085W	Q07500	YDL085W	545	-	Luttik *et al*, 1998
D5 - 1d	*Saccharomyces cerevisiae*	sce:YMR145C	P40215	YMR145C	560	-	Luttik *et al*, 1998
D5 - 3b	*Neurospora crassa*	ncr:NCU08980	Q7S2Y9	NCU08980	577	-	Carneiro *et al*, 2004
D5 - 3c	*Neurospora crassa*	ncr:NCU05225	Q7RVX4	NCU05225	673	-	Melo *et al*, 2001
D5 - 3d	*Neurospora crassa*	ncr:NCU00153	Q7RXQ5	NCU00153	550		Duarte *et al*, 2003
D5 - 14a	*Toxoplasma gondii*	tgo:TGME49_009150	Q3HLX0	TGME49_009150	618	-	Lin *et al*, 2008
D5 - 14b	*Toxoplasma gondii*	tgo:TGME49_088830	S8EWS1	TGME49_088830	657	-	Lin *et al*, 2008
D5 - 15	*Plasmodium falciparum*	pfa:PFI0735c	Q8I302	PFI0735c	533	-	Biagini *et al*, 2006
D5 - 16	*Yarrowia lipolytica*	yli:YALI0F25135g	F2Z699	YALI0F25135g	582	-	Kerscher *et al*, 1999; Eschemann *et al*, 2005
D5 - 17	*Chlamydomonas reinhardtii*	cre:CHLREDRAFT_133334	A8JI60	CHLREDRAFT_133334	615	-	Desplats *et al*, 2009
D5 - 18	*Trypanosoma brucei*	tbr:Tb10.6k15.0960	Q38A31	Tb10.6k15.0960	491	-	Fang and Beattie, 2002

[1] The selected *S. aureus* strain in our dataset does not correspond to the strains used by Schurig-Briccio *et al*, 2014, or Sena *et al*, 2015, but the NDHs-2 present in these strains share 100 % sequence identity.

[2] *A. tumefaciens* species was not present in the KEGG database in September 2014, but the NDH-2 sequence (AGROH133_07261) from *Agrobacterium* spp. has 100 % identity with that referred to in Bernard *et al*, 2006

manually refined in Jalview 2.8.1 [11]. Hydrophobic moments of amphipathic helices when necessary were predicted by HeliQuest [12] (data not shown). Signal peptide recognition was predicted using TargetP 1.1 [13]. Disulfide bonds were predicted using EDBCP [14].

III.1.2.3 Phylogenetic Analysis

Phylogenetic analysis was performed using two different methods, Neighbor-Joining (NJ) and Maximum-likelihood (ML). NJ trees were generated in ClustalX 2.1 [5] with 1000 replicates of bootstrap and 999 generated seeds. ML phylogenetic analysis was performed using RAxML (http://www.trex.uqam.ca/) with 100 replicates of bootstrap and 101 generated seeds [15]. In the first case (NJ tree) we used the 2567 amino acid sequences composing our data set (Figure III.1.1, step 9), and in addition we used the sequences from SQR and FCSD to define an outer group (Figure III.1.1, step 8). In the case of ML tree the data set was reduced to 312 amino acid sequences, due to computational constrains. The 312 amino acid sequences were randomly chosen taking into account the clusters generated by the NJ tree and distribution in the Eukarya, Bacteria or Archaea domains. The generated dendrograms were visualized in Dendroscope v3.2.10 [16].

III.1.2.4 Structural models

Structural models of the proteins clustered in group A and B were predicted by the Protein Homology/analogy Recognition Engine V 2.0 (Phyre2) server [17] (not shown). Protein homology structure

images were analysed using PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.

III.1.3 Results and Discussion

III.1.3.1 NDH-2 is present in genomes from microorganisms from the three domains of life: Eukarya, Bacteria and Archaea

We performed a taxonomic profile of NDH-2s evaluating the presence of their coding genes in all species deposited in the KEGG database (Figure III.1.2 and Supplementary Table III.1.1) [1, 2]. We observed that, from the 1805 species considered in our analysis, 1107 species (corresponding to 61 %) contain at least one gene coding for NDH-2. NDH-2s are present in genomes of microorganisms from all three domains of life, Eukarya, Bacteria and Archaea (Figure III.1.2). In Bacteria, genes coding for NDH-2 are observed in 26 of the 36 phyla and in 60 % of the species. 100 % of the species from Cyanobacteria (40 species), 89 % from Actinobacteria (177 species), 66 % from Proteobacteria superphylum (592 species) and 39 % from Firmicutes (262 species) contain at least one gene encoding NDH-2. In Archaea, we noticed 71 NDH-2s that are present in 32 % of the 143 species, mainly belonging to the *Halobacteriales* (Euryarchaeota) and *Sulfolobales* (Crenarchaeota) orders.

In Fungi, the Ascomycetes and Basidiomycetes phyla have 100 % (72 species) and 96 % (25) of the species with genes encoding NDH-2, respectively, whereas these genes are not observed in the four

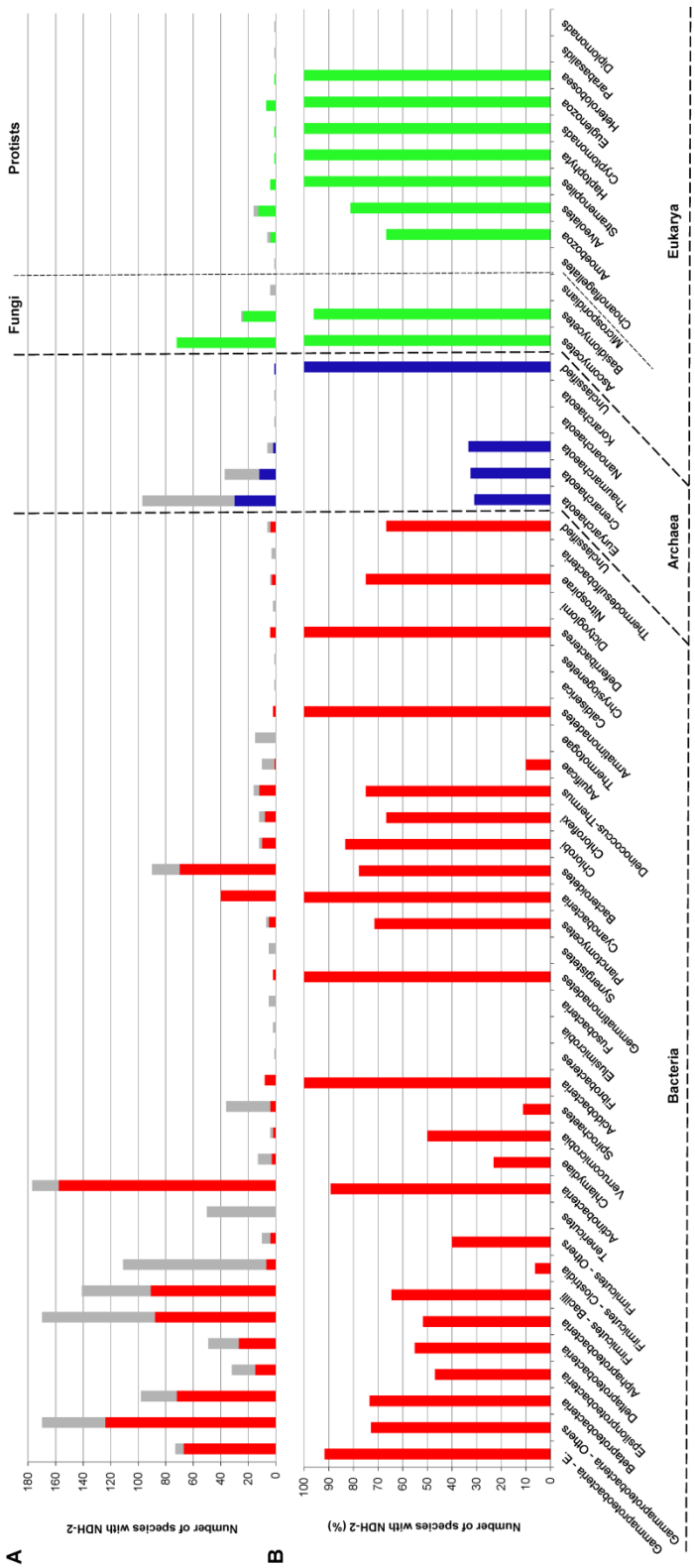

A
B
Number of species with NDH-2
Number of species with NDH-2 (%)
Bacteria
Archaea
Eukarya
Fungi
Protists
Gammaproteobacteria - E.
Gammaproteobacteria - Others
Epsilonproteobacteria
Deltaproteobacteria
Alphaproteobacteria
Firmicutes
Firmicutes - Bacilli
Firmicutes - Clostridia
Firmicutes - Others
Tenericutes
Actinobacteria
Chlamydiae
Verrucomicrobia
Spirochaetes
Acidobacteria
Fibrobacteres
Elusimicrobia
Fusobacteria
Gemmatimonadetes
Synergistetes
Planctomycetes
Cyanobacteria
Bacteroidetes
Chlorobi
Deinococcus-Thermus
Aquificae
Thermotogae
Armatimonadetes
Caldiserica
Chrysiogenetes
Deferribacteres
Dictyoglomi
Nitrospirae
Thermodesulfobacteria
Euryarchaeota
Crenarchaeota
Thaumarchaeota
Nanoarchaeota
Korarchaeota
Unclassified
Ascomycetes
Basidiomycetes
Microsporidians
Chaonoflagellates
Amoebozoa
Alveolates
Stramenopiles
Haptophyta
Cryptomonads
Euglenozoa
Heterokonts
Rhizaria
Diplomonads

Figure III.1.2. Distribution of NDH-2s in the three domains of life. The diagram indicates the number and fraction of species *per* phylum that encode at least one NDH-2. A) The bars in gray indicate the total number of species present in each phylum. These are superimposed by colored bars indicating the number of species *per* phylum with at least one gene encoding a NDH-2. B) Percentage of the species within each phylum containing at least one gene encoding a NDH-2. The color code is: Eukarya, green; Bacteria, red and Archaea, blue. The methods used to retrieve and analyze the NDH-2s are described in Figure III.1.1. Information on the distribution of NDH-2s in animals and plants may be found in Supplementary Figure 3.2.1.

species of the Microsporidia phylum. 80 % of the 39 protist species have at least one gene encoding NDH-2. Genes coding for NDH-2 were not observed in a single species of the Parabasalids and Diplomonads phyla.

NDH-2 seems to be absent in prokaryotes that are strict anaerobes, such as *Tenericutes* (50 species), *Thermotogae* (15 species), *Fusobacteria* (5 species), *Synergistetes* (5 species), *Elusimicrobia* (2 species), *Dictyoglomi* (2 species), *Fibrobacteres* (1 species), *Caldiserica* (1 species), *Chrysiogenetes* (1 species) and *Nanoarchaeota* (1 species).

We observed the simultaneous presence of several genes coding for NDH-2 in many species, in agreement with previous reports. Most of the species from bacterial and archaeal phyla contain on average not more than two copies of NDH-2s, the exceptions being the species from *Actinobacteria* and *Cyanobacteria* phyla with averages of 2.7 and 3.5, respectively. The species from the fungal *Ascomycetes* and *Basidiomycetes* phyla have on average 5.4 and 5.5 copies, respectively,

while those of the phyla from protist kingdom contain on average between 1.2 and 6 NDH-2 homologues. The Species *Thalassiosira pseudonana* from the protist phylum *Stramenopiles* is an illustrative example since it has five NDH-2s, two of which contain recognized mitochondrial targeting signal peptides, and one seems to have a chloroplast targeting signal peptide (Table III.1.3).

The reason for the presence of multiple NDH-2s in the same organism is not clear, but it could be related to their role in different metabolic pathways, for example in respiration or photosynthesis, or/and different organelle localization. Some NDH-2s sequences contain mitochondrial or chloroplast signal peptides and proteins have been detected facing both sides of plastid membranes. Also some of the eukaryotic proteins have an additional sequence insertion, which has been shown in some case to correspond to the EF-hand structural motif (see below and Table III.1.4). This motif is able to bind calcium, suggesting possibilities of protein activity regulation [18, 19].

III.1.3.2 NDH-2 family distribution is not congruent with the taxonomic tree

In order to explore the relations among the different members of the NDH-2 family we generated two dendrograms using two different methods, Neighbour-Joining (NJ) and Maximum-likelihood (ML). The two analyses gave similar results (Figure III.1.3). We observe that close to the root of the dendrogram, the sequences of NDH-2s separate into four distinct groups, which we named A, B, C and D (clockwise direction) in order to facilitate the description of the

Table III.1.3. Signal Peptide prediction for eukaryotic sequences using TargetP 1.1 [13].

		KEGG Category	Species with NDH-2	Number of NDH-2	NDH-2 *per* Species	Mitochondria targeting signal				Chloroplast targeting signal				Secretory pathway targeting signal				Unknown			
						Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species
						no.	%			no.	%			no.	%			no.	%		
Eukarya	Animals	Vertebrates	52	60	1.2	4	8	4	1.0	-	-	-	-	7	13	7	1.0	43	83	49	1.1
		Lancelets	1	2	2.0	-	-	-	-	-	-	-	-	-	-	-	-	1	100	2	2.0
		Ascidians	1	7	7.0	1	100	1	1.0	-	-	-	-	-	-	-	-	1	100	6	6.0
		Echinoderms	1	1	1.0	-	-	-	-	-	-	-	-	-	-	-	-	1	100	1	1.0
		Arthropods	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Nematodes	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Annelids	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Mollusks	1	1	1.0	1	100	1	1.0	-	-	-	-	-	-	-	-	-	-	-	-
		Flatworms	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Cnidarians	2	4	2.0	2	100	3	1.5	-	-	-	-	-	-	-	-	1	50	1	1.0
		Placozoans	1	1	1.0	-	-	-	-	-	-	-	-	-	-	-	-	1	100	1	1.0
		Poriferans	1	2	2.0	1	100	1	1.0	-	-	-	-	1	100	1	1.0	-	-	-	-
	Plants	Eudicots	22	188	8.5	22	100	99	4.5	20	91	33	1.7	12	55	13	1.1	21	95	43	2.0
		Monocots	7	55	7.9	7	100	36	5.1	5	71	5	1.0	2	29	2	1.0	7	100	12	1.7
		Basal Magnoliophyta	1	6	6.0	1	100	2	2.0	1	100	1	1.0	1	100	1	1.0	1	100	2	2.0
		Ferns	1	6	6.0	1	100	1	1.0	1	100	3	3.0	-	-	-	-	1	100	2	2.0
		Mosses	1	7	7.0	1	100	4	4.0	-	-	-	-	1	100	1	1.0	1	100	2	2.0
		Green algae	9	42	4.7	9	100	18	2.0	7	78	11	1.6	3	33	3	1.0	6	67	10	1.7
		Red algae	3	9	3.0	2	67	3	1.5	2	67	3	1.5	-	-	-	-	2	67	3	1.5
	Fungi	Ascomycetes	72	392	5.4	72	100	193	2.7	-	-	-	-	40	56	65	1.6	57	79	134	2.4
		Basidiomycetes	24	133	5.5	24	100	76	3.2	-	-	-	-	1	4	1	1.0	22	92	56	2.5
		Microsporidians	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Protists	Choanoflagellates	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Amoebozoa	4	22	5.5	3	75	5	1.7	-	-	-	-	4	100	5	1.3	4	100	12	3.0
		Alveolates	13	15	1.2	4	31	4	1.0	-	-	-	-	2	15	2	1.0	8	62	9	1.1
		Stramenopiles	4	14	3.5	2	50	3	1.5	1	25	1	1.0	2	50	3	1.5	3	75	7	2.3
		Haptophyta	1	4	4.0	1	100	1	1.0	-	-	-	-	-	-	-	-	1	100	3	3.0
		Cryptomonads	1	6	6.0	-	-	-	-	1	100	1	1.0	1	100	2	2.0	1	100	3	3.0
		Euglenozoa	7	16	2.3	7	100	8	1.1	-	-	-	-	7	100	8	1.1	-	-	-	-
		Heterolobosea	1	5	5.0	1	100	1	1.0	-	-	-	-	1	100	1	1.0	1	100	3	3.0
		Parabasalids	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Diplomonads	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Total	**231**	**998**		**166**	**72**	**464**		**38**	**16**	**58**		**85**	**37**	**115**		**184**	**80**	**361**	

Table III.1.4. Distribution of the putative EF hand motifs.

		KEGG Category	Species with NDH-2	Number of NDH-2	NDH-2 *per* Species	Putative EFhand motif			
						Species with NDH-2		Number of NDH-2	NDH-2 *per* Species
						no.	%		
Eukarya	Animals	Vertebrates	52	60	1.2	-	-	-	-
		Lancelets	1	2	2.0	-	-	-	-
		Ascidians	1	7	7.0	-	-	-	-
		Echinoderms	1	1	1.0	-	-	-	-
		Arthropods	-	-	-	-	-	-	-
		Nematodes	-	-	-	-	-	-	-
		Annelids	-	-	-	-	-	-	-
		Mollusks	1	1	1.0	-	-	-	-
		Flatworms	-	-	-	-	-	-	-
		Cnidarians	2	4	2.0	1	50	1	1.0
		Placozoans	1	1	1.0	-	-	-	-
		Poriferans	1	2	2.0	1	100	1	1.0
	Plants	Eudicots	22	188	8.5	22	100	77	3.5
		Monocots	7	55	7.9	7	100	21	3.0
		Basal Magnoliophyta	1	6	6.0	1	100	1	1.0
		Ferns	1	6	6.0	1	100	1	1.0
		Mosses	1	7	7.0	1	100	3	3.0
		Green algae	9	42	4.7	7	78	10	1.4
		Red algae	3	9	3.0	1	33	1	1.0
	Fungi	Ascomycetes	72	392	5.4	43	60	43	1.0
		Basidiomycetes	24	133	5.5	22	92	22	1.0
		Microsporidians	-	-	-	-	-	-	-
	Protists	Choanoflagellates	-	-	-	-	-	-	-
		Amoebozoa	4	22	5.5	4	100	4	1.0
		Alveolates	13	15	1.2	11	85	11	1.0
		Stramenopiles	4	14	3.5	-	-	-	-
		Haptophyta	1	4	4.0	-	-	-	-
		Cryptomonads	1	6	6.0	-	-	-	-
		Euglenozoa	7	16	2.3	-	-	-	-
		Heterolobosea	1	5	5.0	1	100	1	1.0
		Parabasalids	-	-	-	-	-	-	-
		Diplomonads	-	-	-	-	-	-	-
		Total	**231**	**998**		**123**	**53**	**197**	

dendrogram. These groups do not correspond to those previously proposed by Melo *et al* [20] to classify NDH-2s. In fact, we believe that this previously proposed classification is misleading since the sequences composing its group C are not NDH-2s, but SQRs [6-8]. The color codes of both branches and of the external ring of Figure III.1.3

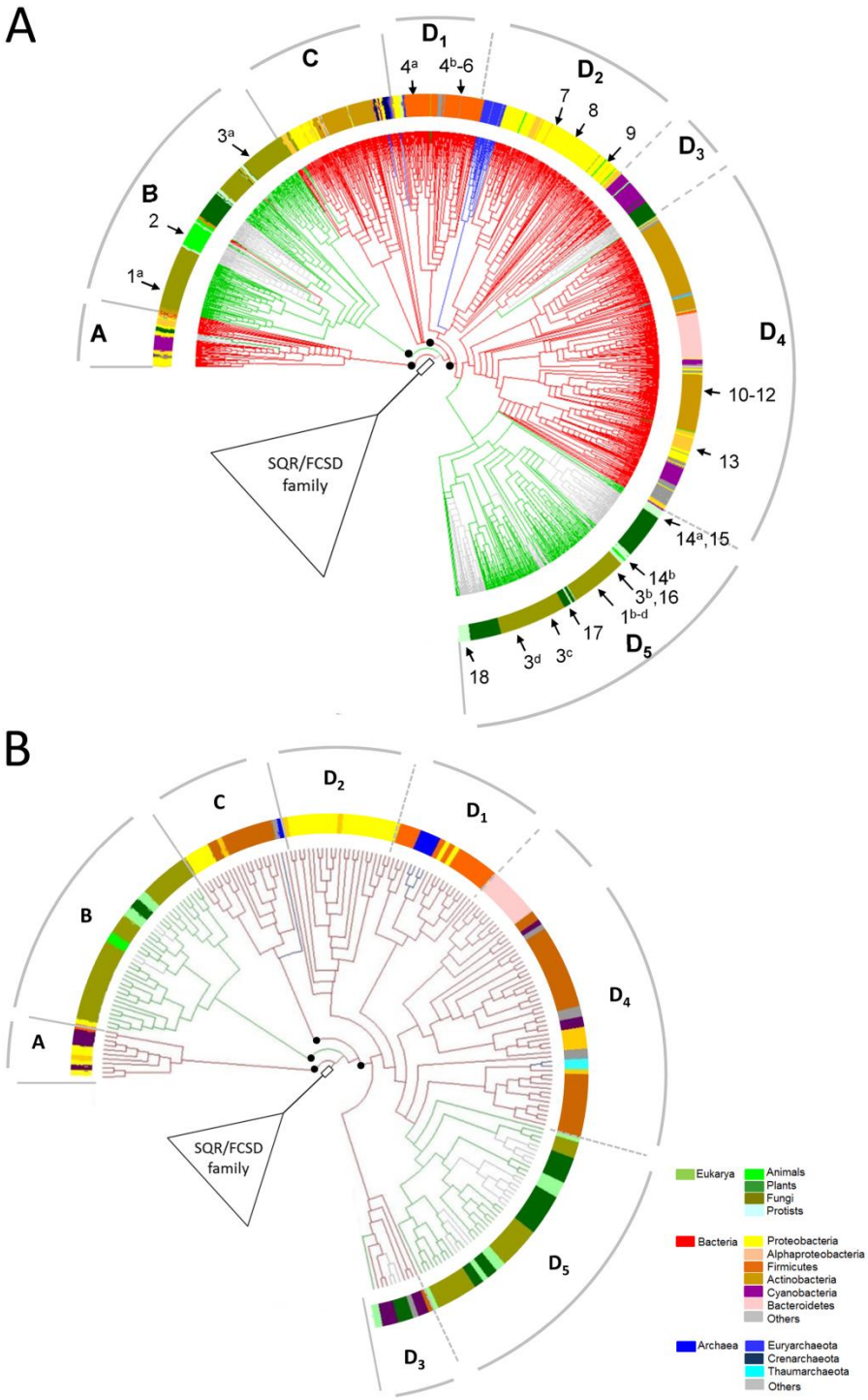

A
C
D1
4a
4b-6
D2
7
8
9
D3
3a
B
2
1a
A
D4
10-12
13
14a,15
14b
3b,16
1b-d
17
D5
18
3d
3c
SQR/FCSD
family
B
D2
C
D1
B
D4
A
SQR/FCSD
family
D5
D3
Eukarya
Animals
Plants
Fungi
Protists
Bacteria
Proteobacteria
Alphaproteobacteria
Firmicutes
Actinobacteria
Cyanobacteria
Bacteroidetes
Others
Archaea
Euryarchaeota
Crenarchaeota
Thaumarchaeota
Others

Figure III.1.3. Dendrogram of NDH-2s. Branches are colored according to the three domains of life from which the organism containing the respective NDH-2 originates: green corresponds to microorganisms from Eukarya (plant and animal sequences are represented in gray), red to bacteria and blue to archaea. The external ring of the dendrogram, in the region of NDH-2s, indicates the phyla containing the organisms with the respective NDH-2 sequences according to the color code indicated in the figure. A) The NJ dendrogram was obtained using 2567 sequences of NDH-2s and sequences from SQRs/FCCDs to define the outer group, midpoint root, 1000 replicates of bootstrap, 999 generated seeds. Sequences of biochemically characterized NDH-2s are indicated as: 1a-d) *S. cerevisiae* [locus tag: YNR074C, YML120C, YDL085W, YMR145C], 2) *H. sapiens* [GenBank: AAH06121], 3a-d) *N. crassa* [NCU09447, NCU05225, NCU00153, NCU08980], 4a-b) *S. aureus* [SA0799, SA0802], 5) *C. thermarum* [EGL84165], 6) *B. pseudofirmus* [BpOF4_04810], 7) *M. capsulatus* [MCA1918], 8) *E. coli* [b1109], 9) *G. oxydans* [GOX1675], 10) *M. tuberculosis* [Rv1854c], 11) *M. smegmatis* [MSMEG_3621], 12) *C. glutamicum* [NCgl1409], 13) *A. tumefaciens* [Ach5_19080], 14a-b) *T. gondii* [TGME49_009150, TGME49_088830], 15) *P. falciparum* [PF07_0085], 16) *Y. lipolytica* [YALI0F25135g], 17) *C. reinhardtii* [CHLREDRAFT_133334], 18) *T. brucei* [Tb10.6k15.0960]. B) The ML dendrogram was obtained using 312 amino acid sequences randomly chosen taking into account the clusters generated by the NJ tree (A) and their distribution in the Eukarya, Bacteria or Archaea domains, midpoint root, 100 replicates of bootstrap and 101 generated seeds.

immediately indicate that these dendrograms are not congruent with the taxonomic tree. For example, sequences from eukaryotic organisms form three groups and those from Archaea are mostly observed in two groups.

Interestingly, we observe that 39 of the sequences present in the first diverging group (group A) have a C-terminal extension of

approximately 200 to 350 amino acid residues. The C-terminal extensions have sequence similarity (>50 %) to proteins of the selenophosphate synthetase family (EC 2.7.9.3) and this observation could indicate a further specialization of this group of enzymes.

The second group to diverge (group B) is composed of eukaryotic sequences including those from fungi and protists. NDH-2 has been considered to be absent from mammals, but our analysis retrieved sequences from 31 mammalian species, including *Homo sapiens*, which contain at least one NDH-2 copy, all included in group B. The protein from *H. sapiens*, AIF-M2 (initially named as apoptosis-inducing factor-homologous mitochondrion-associated inducer of death, AMID [21] and later AIF-M2 [22]), was suggested to be another apoptosis inducing factor (AIF). However this protein is observed to cluster in a group included in the NDH-2 family and not in the group containing the canonic AIF proteins [23], which are also members of the tDBDF superfamily (Figure III.1.4, "8" and "9"). According to our criteria of selection, AIF-M2 is a NDH-2. The possibility of AIF-M2 to function as NDH-2 and its position in the dendrogram (Figure III.1.3) are corroborated by the initial work of Marshall *et al* [21] and a recent report, which showed that the human protein has the NADH:quinone oxidoreductase activity [24]. The nonexistence of NDH-2s in mammals and the fact that these are the only NADH:quinone oxidoreductase expressed in some pathogenic microorganisms were reasons invoked to propose NDH-2 as an attractive target for rational design of specific drugs. The recognition of the existence of these proteins in mammals may undermine this idea.

Besides AIF-M2 from *H. sapiens* (“2” in Figure III.1.3), two other proteins, NDE3 from *N. crassa* [25] and Aif1p from *S. cerevisiae* [26], present in group B, have been studied (Table III.1.2). The so-called AIF homologue from *S. cerevisiae*, Aif1p (Ynr074cp, “1a” in Figure III.1.3), was shown to stimulate apoptotic cell death induced by hydrogen peroxide and to be attached to the inner mitochondrial membrane or alternatively located in the matrix [26]. Interestingly, this protein is different from the yeast Ndi1 [27], mentioned above, which was also suggested to be involved in yeast apoptosis [28]. The sequence of Ndi1 clusters with NDH-2s present in group D (Figure III.1.3, A). The mitochondrial external NDE3 from *N. crassa* (“3a” in Figure III.1.3), one of the four NDH-2s in this fungus, was suggested to be important in specific stages of fungal development [25]. This protein was also observed in the cytosol and was reported to be more closely related to AIF-M2 proteins than to the other NDH-2s [29].

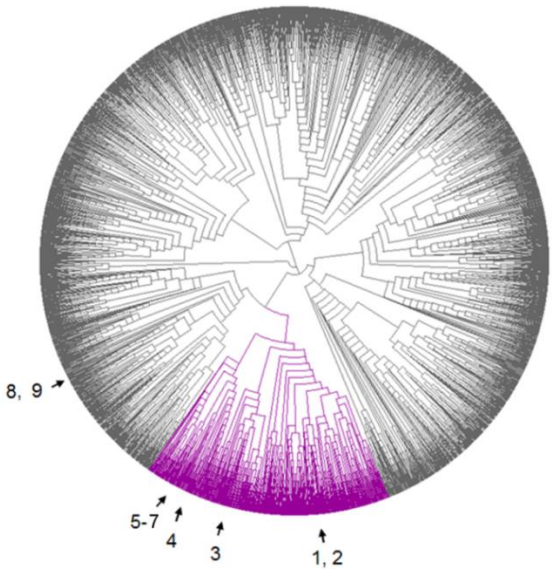


Figure III.1.4. Neighbor-joining dendrogram of representatives from the tDBDF superfamily. The dendrogram was obtained using ~14,000 amino acid sequences, midpoint root, 10 replicates of bootstrap and 3 generated seeds. The sequences highlighted in purple are those designated before as NADH dehydrogenase in which the NDH-2 and sulfide dehydrogenase (SQR/FCSD) families are included. The numbers indicated in the figure refer to proteins with available tertiary structure: 1, Ndi1 from *S. cerevisiae* (YML120C, PDBs: 4G6H and 4G9K)]; 2, NDH-2 from *S. aureus* (SAOUHSC_00878, PDB: 4XDB); 4, FCSD from *A. vinosum* (Alvin_1092, PDB: 1FCD); 5, SQR from *A. ferrooxidans* (Lferr_1469, PDBs: 3SX6, 3SY4 and 3SZF); 6, SQR from *A. aeolicus* (aae:aq_2186, PDBs: 3HYW, 3HYV and 3HYX), 7: SQR from *A. ambivalens* [GI: 238828306, PDBs: 3H8I) (*A. ambivalens* species is not deposited in KEGG´s database. Here is represented by the SQR from *A. hospitalis* (Ahos_0513) which shares 100% identity). The numbers 3, 8 and 9 refer to proteins without tertiary structure available and correspond to AIF-M2 from *H. sapiens* (GenBank: AAH06121), AIF1 from *H. sapiens* (GenBank: NP_001124318)] and AIF3 from *H. sapiens* (Genbank: NP_001018070), respectively.

An interesting feature observed in sequences from group C (specifically in 134 sequences) is the presence of two cysteines residues which we predict to establish a disulfide bond close to the isoalloxazine ring of FAD (Figure III.1.5, A). One of these cysteine residues is located in the proposed quinone binding site [30] and the second in the first β-sheet of the C-terminal domain, in a common position to one of the redox active cysteines of the members of the sulfide dehydrogenase family. Contrary to sulfide dehydrogenase family (SQR and FCSD) that has the NADH pocket blocked (Figure III.1.5, B, in blue), the proteins whose sequences are present in group C seem to have this pocket accessible. This information may suggest

that those 134 sequences have an additional catalytic or regulatory site composed of the above-mentioned disulfide bond or may represent a further specialization step to another yet unknown catalytic activity. None of the enzymes in this group of sequences has been biochemically characterized so far.

The fourth group to diverge (group D) clearly forms several clusters. The first cluster contains almost all sequences from *Firmicutes* and the biochemically characterized enzymes from *S. aureus, Bacillus pseudofirmus* and *C. thermarum* (Table III.1.2). *S. aureus* presents two copies of NDH-2 (tags: SA0799 and SA0802], "4a" and "4b" in Figure III.1.3), which were named NdhF (SA0799, "4a") and NdhC (SA0802, "4b") [31]. Sena *et al* [32] performed a biochemical and biophysical characterization of one of the NDH-2 from *S. aureus* (SA0802, "4b" in Figure III.1.3) and both crystal and solution structures were determined.

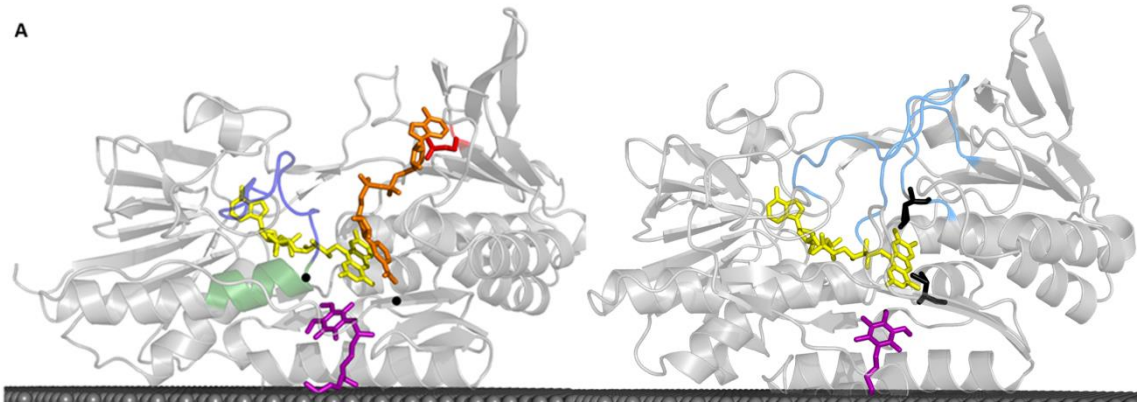


Figure III.1.5. Cartoon representing the structures from NDH-2 and SQR. The structural alignment between these two structures leads an RMSD of 3.2 between the backbone atoms. FAD cofactor (yellow) and quinone (purple) are showed in both

structures, while NADH (orange) is showed in panel A. **A.** NDH-2 crystallographic structure from *S. aureus* [PDB:4XDB]. NADH (orange) and quinone (purple) were co-crystalized with the NDH-2 from *S. cerevisiae* [PDB: 4G6H] and aligned to NDH-2 from *S. aureus* for representation of the NADH and quinone binding pockets. The color code is: blue, EF-hand motif or CxxC motif are localized in this loop (Figure III.1.6, A); red, this sequence position (E201 in *S. aureus*) is suggested to be involved in stabilization of NADH or NADPH (Figure III.1.6, B); green, quinone binding site, localized in the beginning of the corresponding α-helix (Figure III.1.6, C); two black dots, localization of the two cysteine residues found in group C. disulfide bonds were predicted using EDBCP [14] (not shown). **B.** Crystallographic structure of SQR from *Aquifex aeolicus* [PDB:3HYW]. The color code is: black, two redox cysteine residues; blue, NADH pocket is blocked by two loops.

NDH-2 from *C. thermarum* ("5" in Figure III.1.3), also mentioned above, was the first bacterial protein to be structurally characterized and the sequence motif AQxAxQ was proposed to define the quinone binding site [30]. *B. pseudofirmus* has two NDH-2s, but only that designated as ndh-2A [BpOF4_04810] ("6" in Figure III.1.3) was characterized and observed to be able to complement NADH dehydrogenase activity in a deficient *E. coli* strain [33]. The protein referred as Ndh-2B (BpOF4_06965) in the same work was not selected as a NDH-2 according to our criteria. We did a close inspection of the Ndh-2B sequence and observed that it clusters with the disulfide CoA reductase family [3]. In fact, this sequence presents the CxxPY motif typical of the members of the disulfide CoA reductase family. The second NDH-2 of *B. pseudofirmus* (BpOF4_04855) has not been biochemically characterized and clusters with NDH-2 "4a" from *S. aureus*.

The first cluster to diverge in group D also includes NDH-2s from members of the *Deinococcus-Thermus* phylum, but these have not been biochemically studied. Although characterization of two putative NDH-2s from *Thermus thermophilus* has been reported in the literature, NDH-2 (TTC1484, NDH-2A) [34] and a second putative NDH-2 (NrcN) [35], both these sequences do not fulfill our selection criteria. We observe that NDH-2A clusters with proteins belonging to the disulfide CoA reductase family and presents the typical CxxPY motif. NrcN has been described as the catalytic subunit of the multimeric NADH dehydrogenase (NRC complex) expressed in cells grown with nitrate under anoxic conditions [36]. This complex appears to be specific for nitrate respiration of *T. thermophilus*. We notice that NrcN groups with the NADH:ferredoxin oxidoreductase subfamily of the tDBDF superfamily [3].

The enzymes from *Gluconobacter oxydans, Methylococcus capsulatus* and *E. coli*, present in the second cluster, were biochemically characterized (Table III.1.2) [37-39]. Copper reductase activity and/or copper regulation of the activity of NDH-2 from *E. coli* (b1109, “8” in Figure III.1.3) was suggested and a putative CxxC metal binding motif has been identified [40, 41]. We observe the presence of this motif in 199 of the 355 sequences present in this cluster (see below, Figure III.1.6, A). This putative metal binding motif seems to be specific to this cluster since it is not present in the other groups of the dendrogram. NDH-2 from the proteobacterium *G. oxydans* (GOX1675, “9” in Figure III.1.3) was never fully purified but studies of NADH:quinone oxidoreductase activities with membrane vesicles have

been performed and a screening of 304 antimicrobial compounds for new NDH-2 inhibitors was done [37].

NDH-2s from *M. tuberculosis* (Rv1854c) [42, 43], *Mycobacterium smegmatis* (MSMEG_3621 [44]), *Corynebacterium glutamicum* (NCgl1409 [45]) and *Agrobacterium tumefaciens* (Ach5_19080 [46]) (Figure III.1.3 "10" to "13", respectively) were biochemically characterized (Table III.1.2). Mutagenesis studies on the enzyme from *A. tumefaciens* showed that changing a glutamate residue (E203) to a glutamine residue modified the specificity of the enzyme towards NADH or NADPH. NDH-2 from *A. tumefaciens* uses NADH as substrate and the referred mutation increased both the affinity and catalytic efficiency for NADPH [47].

The last cluster comprises exclusively eukaryotic sequences, including 12 proteins from microorganisms, which have been biochemically studied (Table III.1.2). These are three NDH-2s from *S. cerevisiae* ("1b", "1c" and "1d" in Figure III.1.3), three from *N. crassa* ("3b", "3c" and "3d" in Figure III.1.3), two from *T. gondii* ("14a" and "14b" in Figure III.1.3), one from *P. falciparum* ("15" in Figure III.1.3), two from *Yarrowia lipolytica* ("16" in Figure III.1.3) [48, 49], one from *C. reinhardtii* ("17" in Figure III.1.3) and another from *T. brucei* ("18" in Figure III.1.3), some of which already mentioned above.

The *T. gondii* genome encodes two NDH-2s (TGME49_009150 and TGME49_088830, "14a" and "14b" in Figure III.1.3) [50]. The protein TgNDH2-I ("14a" in Figure III.1.3) was expressed in *Y. lipolytica* as an internal enzyme and we observe that it has a sequence insertion where the EF-hand motifs are localized in NDH-2s. Nda2 from *C.*

reinhardtii (CHLREDRAFT_133334, "17" in Figure III.1.3) was produced as a recombinant protein in *E. coli* and the flavin cofactor was shown to be FMN by reverse phase HPLC [51]. It is rather curious to observe the presence of a mononucleotide in a dinucleotide structural fold. Also the two sequences of NDH-2s from the parasite *T. brucei* (Tb927.7.4310 and Tb10.6k15.0960) are present in the dendrogram, one of which (Tb10.6k15.0960, "18" in Figure III.1.3) was investigated and again FMN was suggested to be the flavin prosthetic group [52].

III.1.3.3 Evolutionary considerations

The origin and evolution of NDH-2s are unclear and difficult to track, because it may not always be possible to distinguish them from the other members of the tDBDF superfamily.

Our study shows that NDH-2s form several groups of sequences which are not congruent with the taxonomic tree, suggesting different origins for the eukaryotic sequences and possible lateral gene transfer among prokaryotes. The distribution of NDH-2s from the protists *Dictyostelium discoideum* and *Guillardia theta* are relevant examples. The latter organism contains six genes coding for NDH-2. Three sequences cluster with the first divergent groups of sequences (group A and B), one sequence is included in the group containing those from cyanobacteria and two are present in the group of eukaryotic sequences, which contains Ndi1 from yeast, a classic mitochondrial enzyme. This suggests that the sequence similar to those of cyanobacteria was probably acquired from the chloroplast endosymbiont precursor and those close to classic mitochondrial

enzymes were presumably gained from the mitochondrial endosymbiont precursor. The origin of the third group is less clear, but also a mitochondrial endosymbiont precursor origin may be suggested, which implies that the mitochondrial precursor would carry at least two copies of NDH-2 sequences. The species *T. pseudonana* from the protist phylum Stramenopiles also contains one NDH-2 whose sequence clusters with those from cyanobacteria, in addition to four sequences observed in the last divergent group of eukaryotic sequences. Both *G. theta* and *T. pseudonana* have photosynthetic capacity [53, 54]. Curiously, a similar distribution of NDH-2s from plants is observed (Supplementary Table III.1.1), suggesting likewise three different origins for the respective proteins [19].

NDH-2s from fungi show two origins, which correspond to the bacterial sequences present in the mitochondrial endosymbiont precursor, whereas archaeal NDH-2s seem to include three groups of sequence divergence, all shared with bacterial sequences (Supplementary Table III.1.1). Our data suggest that archaea acquired NDH-2s by lateral gene transfer.

The origin of bacterial NDH-2s is less evident. Bacterial sequences are present in all groups of the dendrogram and even sequences from the same species are sometimes distributed among different groups. Taking this observation into account, as well as the one that only 32 % of archaeal species contain genes coding for NDH-2s and the respective sequences cluster within bacterial sequences, it is tempting to speculate that NDH-2s are of bacterial origin.

III.1.3.4 Properties of NDH-2s

Our results allow systematizing several properties of NDH-2s (Figure III.1.6), including the existence of additional sequence motifs (CxxC and EF-hand like), the distribution of residues possibly responsible for the specificity towards NADH or NADPH (E203 in *A. tumefaciens*, [47]), and the quinone binding motif [30] (Figure III.1.4). We also correlate the presence of NDH-2s with that of other membrane bound NADH:quinone oxidoreductases, namely respiratory Complex I and Na^+-NQR.

We identify the distribution of additional sequence motifs, CxxC and other motifs, assigned in some cases as EF-hand motifs (Figure III.1.5, Figure III.1.6 A and Table III.1.4). The first motif is exclusively present in some proteobacterial sequences included in the second cluster of group D, such as that from *E. coli*. In the case of the second motif, all the 197 sequences containing approximately 70 amino acid residue insertions (which were identified in some cases as an EF motif) are present in the last cluster of group D, but those are not placed together (Figure III.1.6, A). We observe two main clusters, one including fungi proteins and the other containing protist and plant NDH-2s (Table III.1.4).

As mentioned above, the change of a glutamate residue E203 to a glutamine residue modified the specificity of the enzyme from *A. tumefaciens* towards NADH or NADPH [47]. We checked the distribution of the Glu/Asp or Gln/Asn residues possibly responsible for the specificity towards NADH or NADPH, respectively, and observed Glu/Asp in 68 % of the sequences, while Gln/Asn residues

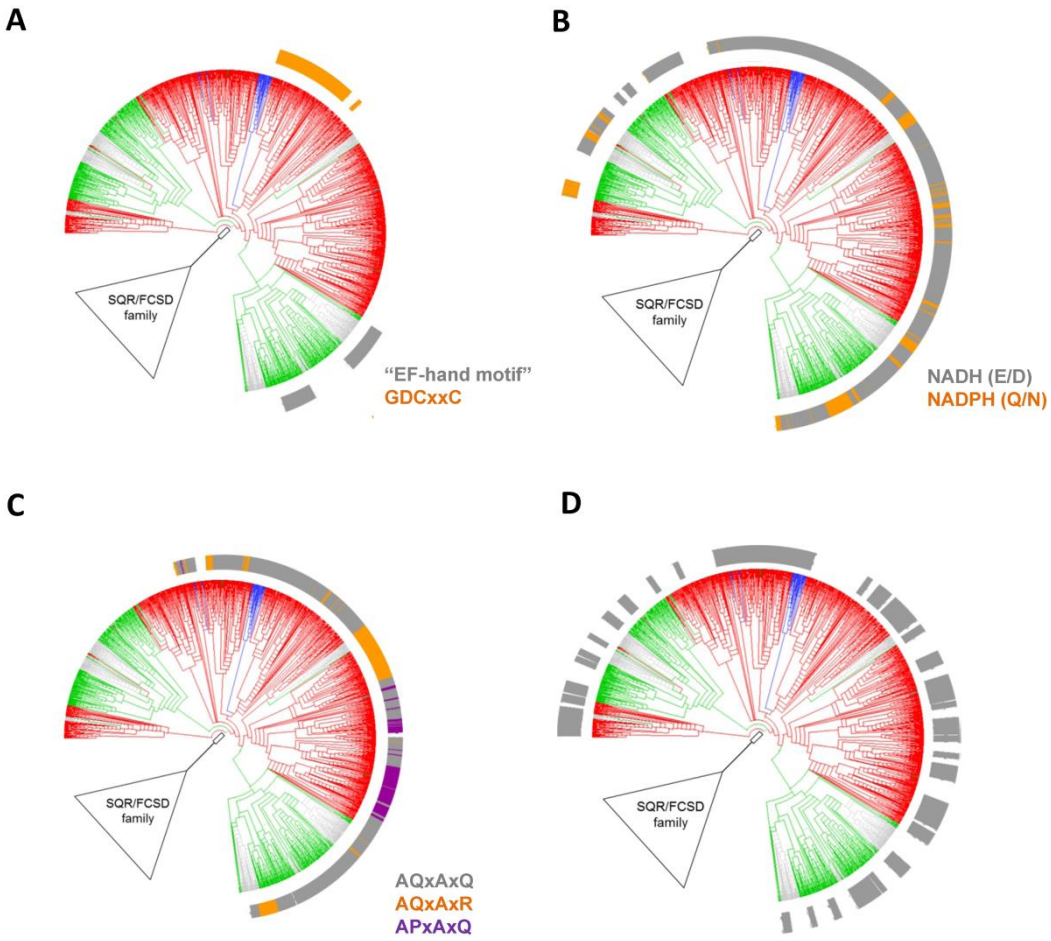


Figure III.1.6. Distribution of several properties of NDH-2s. Neighbor-joining dendrograms as obtained in Fig. III.1.2. The exterior ring indicates: A**)** prediction of the existence of EF-hand like or CxxC motifs; B**)** distribution of the Glu/Asp and Gln/Asn residues that is indicative of the specificity towards NADH or NADPH, respectively; C) distribution of the quinone binding site motifs (AQxAxQ, AQxAxR or APxAxQ); and D) distribution of NDH-2s belonging to microorganisms, which seem not to have all the genes encoding Complex I and/or Na^+-NQR, i.e organisms that seem to have NDH-2 as the only membrane bound NADH:quinone oxidoreductase.

Table III.1.5. Prediction of the quinone binding site motif.

		KEGG Category	Species with NDH-2	Number of NDH-2	AQxAxQ				AQxAxR				APxAxQ			
					Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species
					no.	%			no.	%			no.	%		
Eukarya	Animals	Vertebrates	52	60	2	4	4	2.0	-	-	-	-	1	2	1	1.0
		Lancelets	1	2	-	-	-	-	-	-	-	-	-	-	-	-
		Ascidians	1	7	-	-	-	-	1	100	1	1.0	-	-	-	-
		Echinoderms	1	1	-	-	-	-	-	-	-	-	-	-	-	-
		Arthropods	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Nematodes	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Annelids	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Mollusks	1	1	-	-	-	-	-	-	-	-	-	-	-	-
		Flatworms	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Cnidarians	2	4	1	50	1	1.0	2	100	2	1.0	-	-	-	-
		Placozoans	1	1	-	-	-	-	1	100	1	1.0	-	-	-	-
		Poriferans	1	2	1	100	1	1.0	-	-	-	-	-	-	-	-
	Plants	Eudicots	22	188	22	100	92	4.2	22	100	51	2.3	-	-	-	-
		Monocots	7	55	7	100	27	3.9	7	100	12	1.7	-	-	-	-
		Basal Magnoliophyta	1	6	1	100	2	2.0	1	100	2	2.0	-	-	-	-
		Ferns	1	6	1	100	3	3.0	1	100	2	2.0	-	-	-	-
		Mosses	1	7	1	100	4	4.0	1	100	2	2.0	-	-	-	-
		Green algae	9	42	9	100	25	2.8	-	-	-	-	-	-	-	-
		Red algae	3	9	3	100	6	2.0	-	-	-	-	-	-	-	-
	Fungi	Ascomycetes	72	392	72	100	167	2.3	-	-	-	-	-	-	-	-
		Basidiomycetes	24	133	23	96	57	2.5	-	-	-	-	-	-	-	-
		Microsporidians	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Protists	Choanoflagellates	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Amoebozoa	4	22	4	100	12	3.0	-	-	-	-	-	-	-	-
		Alveolates	13	15	11	85	11	1.0	-	-	-	-	-	-	-	-
		Stramenopiles	4	14	2	50	8	4.0	-	-	-	-	-	-	-	-
		Haptophyta	1	4	1	100	1	1.0	-	-	-	-	-	-	-	-
		Cryptomonads	1	6	1	100	3	3.0	-	-	-	-	-	-	-	-
		Euglenozoa	7	16	-	-	-	-	-	-	-	-	-	-	-	-
		Heterolobosea	1	5	1	100	1	1.0	-	-	-	-	-	-	-	-
		Parabasalids	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Diplomonads	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		Gammaproteobacteria - Enterobacteria	67	81	63	94	63	1.0	-	-	-	-	-	-	-	-
		Gammaproteobacteria - Others	124	181	101	81	104	1.0	9	7	9	1.0	14	11	14	1.0
		Betaproteobacteria	72	115	62	86	85	1.4	5	7	5	1.0	2	3	2	1.0
		Epsilonproteobacteria	15	20	9	60	9	1.0	4	27	4	1.0	-	-	-	-
		Deltaproteobacteria	27	39	1	4	1	1.0	1	4	1	1.0	10	37	13	1.3
		Alphaproteobacteria	88	119	27	31	32	1.2	2	2	2	1.0	38	43	42	1.1
		Firmicutes - Bacilli	91	175	67	74	89	1.3	5	5	5	1.0	-	-	-	-
		Firmicutes - Clostridia	7	8	2	29	2	1.0	-	-	-	-	-	-	-	-
		Firmicutes - Others	4	6	-	-	-	-	-	-	-	-	3	75	4	1.3

	KEGG Category	Species with NDH-2	Number of NDH-2	AQxAxQ				AQxAxR				APxAxQ			
				Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species
				no.	%			no.	%			no.	%		
Bacteria	Tenericutes	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Actinobacteria	158	433	85	54	107	1.3	89	56	146	1.6	2	1	2	1.0
	Chlamydiae	3	4	-	-	-	-	1	33	1	1.0	3	100	3	1.0
	Verrucomicrobia	2	3	-	-	-	-	-	-	-	-	1	50	1	1.0
	Spirochaetes	4	6	-	-	-	-	1	25	1	1.0	1	25	1	1.0
	Acidobacteria	8	12	5	63	6	1.2	1	13	2	2.0	3	38	3	1.0
	Fibrobacteres	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Elusimicrobia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Fusobacteria	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Gemmatimonadetes	2	4	-	-	-	-	1	50	2	2.0	2	100	2	1.0
	Synergistetes	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Planctomycetes	5	6	-	-	-	-	-	-	-	-	4	80	4	1.0
	Cyanobacteria	40	141	39	98	62	1.6	-	-	-	-	24	60	42	1.8
	Bacteroidetes	70	89	54	77	69	1.3	-	-	-	-	9	13	9	1.0
	Chlorobi	10	12	1	10	1	1.0	-	-	-	-	9	90	9	1.0
	Chloroflexi	8	13	-	-	-	-	2	25	2	1.0	6	75	6	1.0
	Deinococcus-Thermus	12	14	8	67	8	1.0	-	-	-	-	4	33	4	1.0
	Aquificae	1	1	1	100	1	1.0	-	-	-	-	-	-	-	-
	Thermotogae	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Armatimonadetes	2	2	-	-	-	-	1	50	1	1.0	1	50	1	1.0
	Caldiserica	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Chrysiogenetes	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Deferribacteres	4	4	-	-	-	-	-	-	-	-	4	100	4	1.0
	Dictyoglomi	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Nitrospirae	3	4	-	-	-	-	2	67	2	1.0	-	-	-	-
	Thermodesulfobacteria	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Unclassified	4	6	-	-	-	-	1	25	1	1.0	-	-	-	-
Archaea	Euryarchaeota	30	51	22	73	26	1.2	7	23	7	1.0	-	-	-	-
	Crenarchaeota	12	16	5	42	5	1.0	-	-	-	-	-	-	-	-
	Thaumarchaeota	2	2	-	-	-	-	-	-	-	-	-	-	-	-
	Nanoarchaeota	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Korarchaeota	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Unclassified	1	2	1	100	1	1.0	-	-	-	-	-	-	-	-
	Total	1107	2567	716	65	1096		168	15	264		141	13	167	
	Eukarya	231	998	163	71	425		36	16	73		1	0	1	
	Bacteria	831	1498	525	63	639		125	15	184		140	17	166	
	Archaea	45	71	28	62	32		7	16	7		-	-	-	

occurred in 10 % (Figure III.1.6, B). This distribution seems to indicate that the majority of the enzymes oxidize NADH rather than NADPH.

The previously proposed quinone binding motif, AQxAxQ (43 %) [30], is a wide-spread property of NDH-2s, and we also recognized the existence of two variations, AQxAxR (10 %) and APxAxQ (7 %) (Figure III.1.5, Figure III.1.6 C and Table III.1.5).

We investigated the relation of the presence of NDH-2s with those of respiratory Complex I and Na^+-NQR (which was further explored in chapter 4). This allowed identifying the organisms in which NDH-2 is the only membrane protein with NADH:quinone oxidoreductase activity (Figure III.1.6, D). We observe that almost all species from protists, *Firmicutes* and *Cyanobacteria*, as well as from all archaeal phyla, have genes coding for NDH-2, but do not contain all the genes coding for other membrane bound NADH:quinone oxidoreductases (Complex I or Na^+-NQR). This observation also extends to some species from fungi, *Actinobacteria* and *Bacteroidetes*.

III.1.4 Final remarks

NDH-2 promotes regeneration of NAD^+ without directly coupling NADH:quinone oxidoreduction to charge translocation, which is different from what is observed for Complex I or Na^+-NQR. In this way, NDH-2 favours NAD^+ regeneration over energy transduction. It is not known what originated first: a system that only oxidized NADH followed by the appearance of a second NADH oxidising system able to promote charge translocation and thus perform energy

transduction or the reverse, a system that already promoted the coupling of NADH oxidation to charge translocation was followed by the appearance of a second, simpler system, that had only the oxidoreduction activity. Considering that energy metabolism could have evolved from a primitive chemiosmotic process, as is intensively discussed [55, 56], we could hypothesize that the first system promoted energy coupling. The diversification of the energetic processes, ensuring delivery of enough energy, would stimulate the development of a simple NAD^+ regenerating system. Thus a possible evolutionary pressure could be a situation in which energy transduction was already assured and NAD^+ availability for cellular processes was the principal need.

Most interestingly, genes coding for NDH-2s are observed in all 40 cyanobacteria species with known complete genomes (the only prokaryotic phylum in which this situation occurs), as well as in 100 % of plant species. This observation inspired us to raise the possibility that NDH-2s appeared with the origin of photosynthesis. Photosynthesis would already contribute extensively to the transmembrane difference of the electrochemical potential and could generate an oversupply of NADH if biosynthetic processes were less needed. This would lead to a scenario with excess of NADH and high membrane potential. In such case, NDH-2 could appear as the solution to regenerate NAD^+ without depending to the membrane potential, i.e., the way to re-oxidize NADH independently of the membrane potential since NDH-2 contribute indirectly to it by reducing quinones. This idea is totally speculative, but we believe it will stimulate further

discussion on the existence and evolution of alternative metabolic pathways and enzymes, which bring robustness and also versatility to biological systems.

III.1.5 Acknowledgments

We acknowledge M. Teixeira, R. O. Louro, T. Catarino and P. Refojo for the critical reading of the manuscript. FVS is recipient of a fellowship by Fundação para a Ciência e a Tecnologia (PD/BD/113985/2015). The work was funded by Fundação para a Ciência e a Tecnologia (PTDC/BBB-BQB/2294/2012 to M.M.P). ITQB is supported by Fundação para a Ciência e a Tecnologia through R&D Unit, UID/CBQ/04612/2013.

III.1.6 References

1. Kanehisa, M. & Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res.* 28, 27-30.

2. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation, *Nucleic Acids Res.* 44, D457-62.

3. Ojha, S., Meng, E. C. & Babbitt, P. C. (2007) Evolution of function in the "two dinucleotide binding domains" flavoproteins, *PLoS computational biology.* 3, e121.

4. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences, *Bioinformatics.* 26, 680-2.

5. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D.,

Gibson, T. J. & Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics.* 23, 2947-8.

6. Cherney, M. M., Zhang, Y., Solomonson, M., Weiner, J. H. & James, M. N. (2010) Crystal structure of sulfide:quinone oxidoreductase from Acidithiobacillus ferrooxidans: insights into sulfidotrophic respiration and detoxification, *Journal of molecular biology.* 398, 292-305.

7. Marcia, M., Ermler, U., Peng, G. & Michel, H. (2009) The structure of Aquifex aeolicus sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration, *Proceedings of the National Academy of Sciences of the United States of America.* 106, 9625-30.

8. Brito, J. A., Sousa, F. L., Stelter, M., Bandeiras, T. M., Vonrhein, C., Teixeira, M., Pereira, M. M. & Archer, M. (2009) Structural and functional insights into sulfide:quinone oxidoreductase, *Biochemistry.* 48, 5613-22.

9. Chen, Z. W., Koh, M., Van Driessche, G., Van Beeumen, J. J., Bartsch, R. G., Meyer, T. E., Cusanovich, M. A. & Mathews, F. S. (1994) The structure of flavocytochrome c sulfide dehydrogenase from a purple phototrophic bacterium, *Science.* 266, 430-2.

10. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Res.* 30, 3059-66.

11. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench, *Bioinformatics.* 25, 1189-91.

12. Gautier, R., Douguet, D., Antonny, B. & Drin, G. (2008) HELIQUEST: a web server to screen sequences with specific alpha-helical properties, *Bioinformatics.* 24, 2101-2.

13. Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence, *Journal of molecular biology.* 300, 1005-16.

14. Lippi, M., Passerini, A., Punta, M., Rost, B. & Frasconi, P. (2008) MetalDetector: a web server for predicting metal-binding sites and disulfide bridges in proteins from sequence, *Bioinformatics.* 24, 2094-5.

15. Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models, *Bioinformatics.* 22, 2688-90.

16. Huson, D. H. & Scornavacca, C. (2012) Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks, *Systematic biology.* 61, 1061-7.

17. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nature protocols.* 10, 845-58.

18. Melo, A. M., Duarte, M., Moller, I. M., Prokisch, H., Dolan, P. L., Pinto, L., Nelson, M. A. & Videira, A. (2001) The external calcium-dependent NADPH dehydrogenase from Neurospora crassa mitochondria, *The Journal of biological chemistry.* 276, 3947-51.

19. Michalecka, A. M., Svensson, A. S., Johansson, F. I., Agius, S. C., Johanson, U., Brennicke, A., Binder, S. & Rasmusson, A. G. (2003) Arabidopsis genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light, *Plant physiology.* 133, 642-52.

20. Melo, A. M., Bandeiras, T. M. & Teixeira, M. (2004) New insights into type II NAD(P)H:quinone oxidoreductases, *Microbiology and molecular biology reviews : MMBR.* 68, 603-16.

21. Marshall, K. R., Gong, M., Wodke, L., Lamb, J. H., Jones, D. J. L., Farmer, P. B., Scrutton, N. S. & Munro, A. W. (2005) The human apoptosis-inducing protein AMID is an oxidoreductase with a modified flavin cofactor and DNA binding activity, *The Journal of biological chemistry.* 280, 30735-30740.

22. Gong, M., Hay, S., Marshall, K. R., Munro, A. W. & Scrutton, N. S. (2007) DNA binding suppresses human AIF-M2 activity and provides a connection between redox chemistry, reactive oxygen species, and apoptosis, *The Journal of biological chemistry.* 282, 30331-40.

23. Sevrioukova, I. F. (2011) Apoptosis-inducing factor: structure, function, and redox regulation, *Antioxidants & redox signaling.* 14, 2545-79.

24. Elguindy, M. M. & Nakamaru-Ogiso, E. (2015) Apoptosis Inducing Factor (AIF) and its Family Member, AMID, are Rotenone-Sensitive NADH:ubiquinone Oxidoreductases (NDH-2), *The Journal of biological chemistry.* 290, 20815-20826.

25. Carneiro, P., Duarte, M. & Videira, A. (2007) The external alternative NAD(P)H dehydrogenase NDE3 is localized both in the mitochondria and in the cytoplasm of *Neurospora crassa*, *Journal of molecular biology.* 368, 1114-21.

26. Wissing, S., Ludovico, P., Herker, E., Buttner, S., Engelhardt, S. M., Decker, T., Link, A., Proksch, A., Rodrigues, F., Corte-Real, M., Frohlich, K. U., Manns, J., Cande, C., Sigrist, S. J., Kroemer, G. & Madeo, F. (2004) An AIF orthologue regulates apoptosis in yeast, *The Journal of cell biology.* 166, 969-74.

27. Velazquez, I. & Pardo, J. P. (2001) Kinetic characterization of the rotenone-insensitive internal NADH: ubiquinone oxidoreductase of mitochondria from *Saccharomyces cerevisiae*, *Archives of biochemistry and biophysics.* 389, 7-14.

28. Cui, Y., Zhao, S., Wu, Z., Dai, P. & Zhou, B. (2012) Mitochondrial release of the NADH dehydrogenase Ndi1 induces apoptosis in yeast, *Molecular biology of the cell.* 23, 4373-82.

29. Carneiro, P., Duarte, M. & Videira, A. (2012) Characterization of apoptosis-related oxidoreductases from *Neurospora crassa*, *PloS one.* 7, e34270.

30. Heikal, A., Nakatani, Y., Dunn, E., Weimar, M. R., Day, C. L., Baker, E. N., Lott, J. S., Sazanov, L. A. & Cook, G. M. (2014) Structure of the bacterial type II NADH dehydrogenase: a monotopic membrane protein with an essential role in energy generation, *Molecular microbiology.* 91, 950-64.

31. Schurig-Briccio, L. A., Yano, T., Rubin, H. & Gennis, R. B. (2014) Characterization of the type 2 NADH:menaquinone oxidoreductases from *Staphylococcus aureus* and the bactericidal action of phenothiazines, *Biochimica et biophysica acta.* 1837, 954-63.

32. Sena, F. V., Batista, A. P., Catarino, T., Brito, J. A., Archer, M., Viertler, M., Madl, T., Cabrita, E. J. & Pereira, M. M. (2015) Type-II NADH:quinone oxidoreductase from Staphylococcus aureus has two distinct binding sites and is rate limited by quinone reduction, *Molecular microbiology.* 98, 272-288.

33. Liu, J., Krulwich, T.A., Hicks, D.B. (2008) Purification of two putative type II NADH dehydrogenases with different substrate specificities from alkaliphilic *Bacillus pseudofirmus* OF4, *Biochimica et biophysica acta.* 1777, 453-461.

34. Yagi, T., Hon-nami, K. & Ohnishi, T. (1988) Purification and characterization of two types of NADH-quinone reductase from *Thermus thermophilus* HB-8, *Biochemistry.* 27, 2008-13.

35. Venkatakrishnan, P., Lencina, A. M., Schurig-Briccio, L. A. & Gennis, R. B. (2013) Alternate pathways for NADH oxidation in *Thermus thermophilus* using type 2 NADH dehydrogenases, *Biological chemistry.* 394, 667-76.

36. Cava, F., Zafra, O., Magalon, A., Blasco, F. & Berenguer, J. (2004) A new type of NADH dehydrogenase specific for nitrate respiration in the extreme thermophile *Thermus thermophilus*, *The Journal of biological chemistry.* 279, 45369-78.

37. Mogi, T., Matsushita, K., Murase, Y., Kawahara, K., Miyoshi, H., Ui, H., Shiomi, K., Omura, S. & Kita, K. (2009) Identification of new inhibitors for alternative NADH dehydrogenase (NDH-II), *Fems Microbiol Lett.* 291, 157-161.

38. Cook, S. A. & Shiemke, A. K. (2002) Evidence that a type-2 NADH : quinone oxidoreductase mediates electron transfer to particulate methane monooxygenase in *Methylococcus capsulatus*, *Archives of biochemistry and biophysics.* 398, 32-40.

39. Bjorklof, K., Zickermann, V. & Finel, M. (2000) Purification of the 45 kDa, membrane bound NADH dehydrogenase of *Escherichia coli* (NDH-2) and analysis of its interaction with ubiquinone analogues, *Febs Lett.* 467, 105-110.

40. Rapisarda, V. A., Montelongo, L. R., Farias, R. N. & Massa, E. M. (1999) Characterization of an NADH-linked cupric reductase activity from the *Escherichia coli* respiratory chain, *Archives of biochemistry and biophysics.* 370, 143-50.

41. Rapisarda, V. A., Chehin, R. N., De Las Rivas, J., Rodriguez-Montelongo, L., Farias, R. N. & Massa, E. M. (2002) Evidence for Cu(I)-thiolate ligation and prediction of a putative copper-binding site in the *Escherichia coli* NADH dehydrogenase-2, *Archives of biochemistry and biophysics.* 405, 87-94.

42. Yano, T., Li, L. S., Weinstein, E., Teh, J. S. & Rubin, H. (2006) Steady-state kinetics and inhibitory action of antitubercular phenothiazines on *Mycobacterium tuberculosis* type-II NADH-menaquinone oxidoreductase (NDH-2), *Journal of molecular biology.* 281, 11456-11463.

43. Yano, T., Rahimian, M., Aneja, K. K., Schechter, N. M., Rubin, H. & Scott, C. P. (2014) *Mycobacterium tuberculosis* Type II NADH-Menaquinone Oxidoreductase

Catalyzes Electron Transfer through a Two-Site Ping-Pong Mechanism and Has Two Quinone-Binding Sites, *Biochemistry.* 53, 1179-1190.

44. Vilcheze, C., Weisbrod, T. R., Chen, B., Kremer, L., Hazbon, M. H., Wang, F., Alland, D., Sacchettini, J. C. & Jacobs, W. R., Jr. (2005) Altered NADH/NAD+ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria, *Antimicrobial agents and chemotherapy.* 49, 708-20.

45. Nantapong, N., Otofuji, A., Migita, C. T., Adachi, O., Toyama, H. & Matsushita, K. (2005) Electron transfer ability from NADH to menaquinone and from NADPH to oxygen of type IINADH dehydrogenase of *Corynebacterium glutamicum*, *Biosci Biotech Bioch.* 69, 149-159.

46. Bernard, L., Desplats, C., Mus, F., Cuine, S., Cournac, L. & Peltier, G. (2006) *Agrobacterium tumefaciens* type II NADH dehydrogenase. Characterization and interactions with bacterial and thylakoid membranes, *The FEBS journal.* 273, 3625-37.

47. Desplats, C., Beyly, A., Cuine, S., Bernard, L., Cournac, L. & Peltier, G. (2007) Modification of substrate specificity in single point mutants of *Agrobacterium tumefaciens* type II NADH dehydrogenase, *Febs Lett.* 581, 4017-22.

48. Kerscher, S. J., Okun, J. G. & Brandt, U. (1999) A single external enzyme confers alternative NADH:ubiquinone oxidoreductase activity in *Yarrowia lipolytica*, *Journal of cell science.* 112 (Pt 14), 2347-54.

49. Eschemann, A., Galkin, A., Oettmeier, W., Brandt, U. & Kerscher, S. (2005) HDQ (1-hydroxy-2-dodecyl-4(1H)quinolone), a high affinity inhibitor for mitochondrial alternative NADH dehydrogenase: evidence for a ping-pong mechanism, *The Journal of biological chemistry.* 280, 3138-42.

50. Lin, S. S., Kerscher, S., Saleh, A., Brandt, U., Gross, U. & Bohne, W. (2008) The *Toxoplasma gondii* type-II NADH dehydrogenase TgNDH2-I is inhibited by 1-hydroxy-2-alkyl-4(1H)quinolones, *Biochimica et biophysica acta.* 1777, 1455-62.

51. Desplats, C., Mus, F., Cuine, S., Billon, E., Cournac, L. & Peltier, G. (2009) Characterization of Nda2, a plastoquinone-reducing type II NAD(P)H dehydrogenase in chlamydomonas chloroplasts, *The Journal of biological chemistry.* 284, 4148-57.

52. Fang, J. & Beattie, D. S. (2002) Novel FMN-containing rotenone-insensitive NADH dehydrogenase from *Trypanosoma brucei* mitochondria: isolation and characterization, *Biochemistry.* 41, 3065-72.

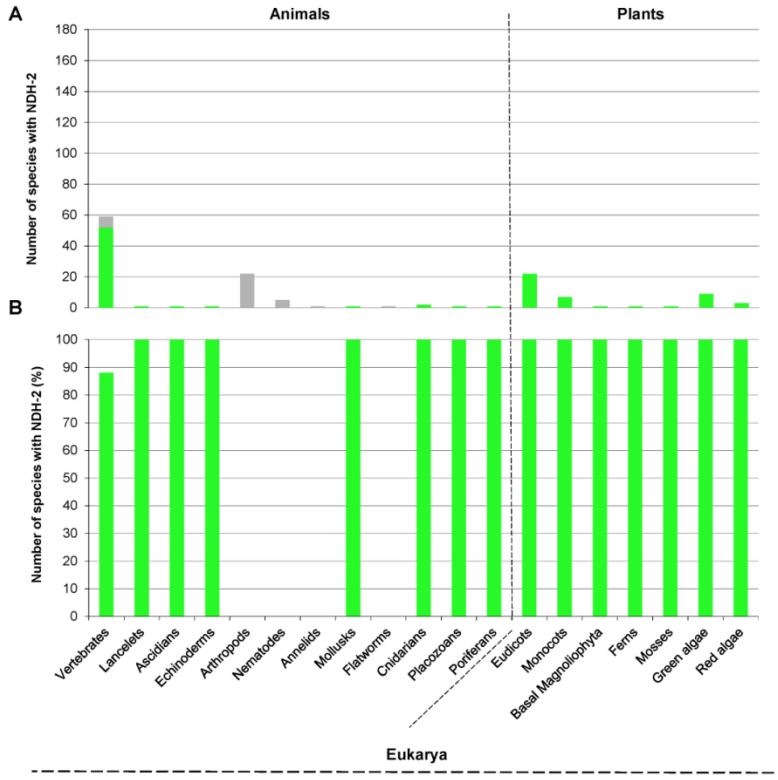
53. Armbrust, E. V., Berges, J. A., Bowler, C., Green, B. R., Martinez, D., Putnam, N. H., Zhou, S., Allen, A. E., Apt, K. E., Bechner, M., Brzezinski, M. A., Chaal, B. K., Chiovitti, A., Davis, A. K., Demarest, M. S., Detter, J. C., Glavina, T., Goodstein, D., Hadi, M. Z., Hellsten, U., Hildebrand, M., Jenkins, B. D., Jurka, J., Kapitonov, V. V., Kroger, N., Lau, W. W., Lane, T. W., Larimer, F. W., Lippmeier, J. C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M. S., Palenik, B., Pazour, G. J., Richardson, P. M., Rynearson, T. A., Saito, M. A., Schwartz, D. C., Thamatrakoln, K., Valentin, K., Vardi, A., Wilkerson, F. P. & Rokhsar, D. S. (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism, *Science.* 306, 79-86.

54. Douglas, S. E. & Penny, S. L. (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae, *Journal of molecular evolution.* 48, 236-44.

55. Lane, N. & Martin, W. F. (2012) The origin of membrane bioenergetics, *Cell.* 151, 1406-16.

56. Nitschke, W. & Russell, M. J. (2009) Hydrothermal focusing of chemical and chemiosmotic energy, supported by delivery of catalytic Fe, Ni, Mo/W, Co, S and Se, forced life to emerge, *Journal of molecular evolution.* 69, 481-96.

III.1.7 Supplementary Material



Supplementary Figure III.1.1. Distribution of NDH-2s in the Animals and Plants kingdoms. Diagram indicating the number of species *per* phylum in the Animals and Plants Kingdoms coding at least one NDH-2. **A.** The bars in gray indicate the total number of species present in each phylum. These are superimposed by colored bars indicating the number of species *per* phylum with at least one gene encoding a NDH-2. **B.** Percentage of the species within each phylum containing at least one gene encoding a NDH-2. The color code is: Eukarya, green.

Supplementary Table III.1.1. Taxonomic profile of NDH-2 family in the three domains of life

KEGG Category			Species with NDH-2				Species with NDH-2 in group (%)							
KEGG Category	Class	Species	Species with NDH-2 no.	Species with NDH-2 %	Number of NDH-2	NDH-2 *per* Species	A	B	C	D1	D2	D3	D4	D5
Vertebrates	6	59	52	88	60	1.2	-	100	-	-	4	-	2	-
Lancelets	1	1	1	100	2	2.0	-	100	-	-	-	-	-	-
Ascidians	1	1	1	100	7	7.0	-	100	-	-	-	-	-	100
Echinoderms	1	1	1	100	1	1.0	-	100	-	-	-	-	-	-
Arthropods	2	22	0	-	-	-	-	-	-	-	-	-	-	-
Nematodes	1	5	0	-	-	-	-	-	-	-	-	-	-	-
Annelids	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Mollusks	1	1	1	100	1	1.0	-	-	-	-	-	-	-	100
Flatworms	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Cnidarians	1	2	2	100	4	2.0	-	50	-	-	-	-	-	100
Placozoans	1	1	1	100	1	1.0	-	-	-	-	-	-	-	100
Poriferans	1	1	1	100	2	2.0	-	-	-	-	-	-	-	100
Eudicots	10	22	22	100	188	8.5	-	95	-	-	-	91	-	100
Monocots	2	7	7	100	55	7.9	-	100	-	-	-	86	-	100
Basal Magnoliophyta	1	1	1	100	6	6.0	-	100	-	-	-	100	-	100
Ferns	1	1	1	100	6	6.0	-	100	-	-	-	100	-	100
Mosses	1	1	1	100	7	7.0	-	100	-	-	-	100	-	100
Green algae	1	9	9	100	42	4.7	67	44	-	-	-	89	-	100
Red algae	1	3	3	100	9	3.0	67	33	-	-	-	100	-	67
Ascomycetes	7	72	72	100	392	5.4	-	81	10	-	-	-	-	100
Basidiomycetes	1	25	24	96	133	5.5	-	92	-	-	-	-	-	100
Microsporidians	1	4	0	-	-	-	-	-	-	-	-	-	-	-
Choanoflagellates	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Amoebozoa	3	6	4	67	22	5.5	-	100	-	-	-	-	-	100
Alveolates	2	16	13	81	15	1.2	-	-	15	-	-	-	-	85
Stramenopiles	3	4	4	100	14	3.5	-	25	-	-	-	25	-	100
Haptophyta	1	1	1	100	4	4.0	-	-	-	-	-	-	-	100
Cryptomonads	1	1	1	100	6	6.0	100	100	-	-	-	100	-	100
Euglenozoa	1	7	7	100	16	2.3	-	100	-	-	-	-	-	100
Heterolobosea	1	1	1	100	5	5.0	-	100	-	-	-	-	-	100
Parabasalids	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Diplomonads	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Gammaproteobacteria - Enterobacteria	35	73	67	92	81	1.2	-	-	19	-	100	-	-	-
Gammaproteobacteria - Others	72	170	124	73	181	1.5	13	-	17	-	86	-	12	-
Betaproteobacteria	56	98	72	73	115	1.6	8	-	18	-	86	-	10	-
Epsilonproteobacteria	10	32	15	47	20	1.3	-	-	53	47	-	-	-	-
Deltaproteobacteria	28	49	27	55	39	1.4	48	-	7	-	-	-	59	-
Alphaproteobacteria	82	170	88	52	119	1.4	23	1	18	-	26	-	51	-
Firmicutes - Bacilli	32	141	91	65	175	1.9	3	-	2	100	-	-	-	-
Firmicutes - Clostridia	54	111	7	6	8	1.1	14	-	29	57	-	-	-	-
Firmicutes - Others	8	10	4	40	6	1.5	-	-	-	-	-	-	100	-
Tenericutes	8	50	0	-	-	-	-	-	-	-	-	-	-	-
Actinobacteria	77	177	158	89	433	2.7	7	5	44	1	-	-	99	-

KEGG Category	Class	Species	Species with NDH-2		Number of NDH-2	NDH-2 *per* Species	Species with NDH-2 in group (%)							
			no.	%			A	B	C	D1	D2	D3	D4	D5
Chlamydiae	6	13	3	23	4	1.3	-	-	-	-	-	-	100	-
Verrucomicrobia	4	4	2	50	3	1.5	-	-	-	-	-	-	100	-
Spirochaetes	7	36	4	11	6	1.5	-	-	50	-	-	-	100	-
Acidobacteria	6	8	8	100	12	1.5	-	-	-	-	-	13	100	-
Fibrobacteres	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Elusimicrobia	2	2	0	-	-	-	-	-	-	-	-	-	-	-
Fusobacteria	5	5	0	-	-	-	-	-	-	-	-	-	-	-
Gemmatimonadetes	2	2	2	100	4	2.0	-	-	-	-	-	-	100	-
Synergistetes	5	5	0	-	-	-	-	-	-	-	-	-	-	-
Planctomycetes	6	7	5	71	6	1.2	20	-	-	-	-	-	100	-
Cyanobacteria	32	40	40	100	141	3.5	68	-	8	-	-	98	73	-
Bacteroidetes	63	90	70	78	89	1.3	1	-	1	-	-	-	100	-
Chlorobi	7	12	10	83	12	1.2	-	-	-	-	-	-	100	-
Chloroflexi	9	12	8	67	13	1.6	-	-	13	38	-	-	100	-
Deinococcus-Thermus	6	16	12	75	14	1.2	17	-	-	58	-	-	42	-
Aquificae	8	10	1	10	1	1.0	-	-	100	-	-	-	-	-
Thermotogae	7	15	0	-	-	-	-	-	-	-	-	-	-	-
Armatimonadetes	2	2	2	100	2	1.0	-	-	-	-	-	-	100	-
Caldiserica	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Chrysiogenetes	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Deferribacteres	4	4	4	100	4	1.0	-	-	50	-	-	-	50	-
Dictyoglomi	1	2	0	-	-	-	-	-	-	-	-	-	-	-
Nitrospirae	3	4	3	75	4	1.3	-	-	33	-	-	-	100	-
Thermodesulfobacteria	2	3	0	-	-	-	-	-	-	-	-	-	-	-
Unclassified	4	6	4	67	6	1.5	-	-	50	-	-	-	50	-
Euryarchaeota	56	97	30	31	51	1.7	-	-	17	-	83	-	-	-
Crenarchaeota	20	37	12	32	16	1.3	-	-	100	-	-	-	-	-
Thaumarchaeota	5	6	2	33	2	1.0	-	-	-	-	-	-	100	-
Nanoarchaeota	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Korarchaeota	1	1	0	-	-	-	-	-	-	-	-	-	-	-
Unclassified	1	1	1	100	2	2.0	-	-	-	-	100	-	-	-
Total	**799**	**1805**	**1107**	**61**	**2567**									
Eukarya	**59**	**280**	**231**	**83**	**998**									
Bacteria	**656**	**1382**	**831**	**60**	**1498**									
Archaea	**84**	**143**	**45**	**32**	**71**									

III.2 - NDH-2 family: Identification of conserved structural and functional elements involved in the catalytic mechanism

This section is based on the following manuscript for publication:

Marreiros BC, Batista AP, Sena FV, Sousa FM, Oliveira SA, Soares, CM, Pereira MM "Type II NADH:quinone oxidoreductase family: Identification of conserved structural and functional elements involved in the catalytic mechanism"

Author contribution:

B.C.M. participated in the design of the study, performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

III.1 - NDH-2 family: Identification of conserved structural and functional elements involved in the catalytic mechanism

III.2.1 Summary

Type II NADH:quinone oxidoreductases (NDH-2s) are membrane proteins involved in respiratory chains. These proteins contribute indirectly to the establishment of the transmembrane difference of electrochemical potential by catalyzing the reduction of quinone by oxidation of NAD(P)H. NDH-2s are wide spread enzymes being present in the three domains of life.

In this work, we explored the catalytic mechanism of NDH-2 by investigating the common elements of all NDH-2s, based on the rational that conservation of such elements reflects their structural/functional importance. We identified conserved sequence motifs and structural elements among 1762 NDH-2s. We recognized the presence of two proton pathways possibly involved in the protonation of the quinone. Our results led us to propose the first catalytic mechanism for NDH-2 family, where a conserved glutamate residue, E172 (*Staphylococcus aureus*) plays a determinant role in NADH binding and proton transfer to the quinone pocket. This catalytic mechanism may also provide further insights into that of the other members of the two-Dinucleotide Binding Domains Flavoproteins (tDBDF) family.

III.2.2 Material and Methods

III.2.2.1 Sequence analysis

We have previously used KEGG data base to select the members of the NDH-2 family (2567 NDH-2s) and performed their taxonomic analysis. We observed that NDH-2 family is distributed in four main branches which we called groups A to D [1].

Identity network analysis performed showed that groups A and B share less than 33 % identity with the main groups C to D. Group B is mainly composed of eukaryotic enzymes that do not present the typical quinone binding site (AQxAxQ), or any of its alternatives (AQxAxR and APxAxQ). We noticed that ~62 % of the NDH-2s present in group C may have a disulphide bond near the proposed quinone binding site and that these sequences do not have a typical quinone binding site [1]. In this analysis we opted to exclude groups A to C from the amino acid conservation analysis. In this way, we aligned the remaining 1779 NDH-2s sequences (~70 % amino acid sequences from the NDH-2 family) using PROMALS3D [2]. We manually refined our data set using Jalview 2.8.1 [3] for which we took into account three criteria: (a) existence of two GxGxxG like motifs for interaction with FAD and NAD(P)H and included few variations, namely GxGxxA motif; (b) presence of a C-terminal amino acid extension for membrane interaction (C-terminal domain), and (c) absence of other domains fused at the N- or C-terminal. Our final data set included, in this way, 1762 amino acid sequences (distributed in the three domains of life,

Eukarya, Bacteria and Archaea). Covariance between amino acid residues in NDH-2 family were predicted by using MISTIC [4].

III.2.2.2. Secondary and tertiary structures analysis

The crystallographic structures used as templates were those from *S. aureus* (PDB:4XDB, [5]), *Caldalkalibacillus thermarum* (PDB:4NWZ, [6]) and *Saccharomyces cerevisiae* (PDB:4G73 [7]). Images of the structures were generated using PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC. Secondary structure of NDH-2 from *S. aureus* was visualized using Stride [8]. All distances measurements presented below were performed between the closest hydrogen atoms of both objects and should be considered as an approximate value.

III.2.2.3 Theoretical calculations

In order to locate the groups likely to be involved in proton transfer, we have determined the pH titration curves for all the protonable residues in NDH-2 from *S. aureus* (pdb code: 4XDB [5]) and from *S. cerevisiae* (pdb code: 4G73 [7]) using methodologies for studying the thermodynamics of proton binding described in detail in [9, 10]. These methodologies use a combination of Poisson-Boltzmann (PB) calculations, performed with the program MEAD (version 2.2.9) [11-13], and Metropolis Monte Carlo (MC) simulations, using the program PETIT (version 1.5) [10]. For the *S. aureus* enzyme, the PB/MC calculations were done with the flavin adenine dinucleotide cofactor in two fixed redox states: the fully oxidized (FAD) and the fully reduced

($FADH_2$) states. For the *S. cerevisiae* enzyme, three systems were simulated: the protein with a FAD, the protein with $FADH_2$ and the protein with the $FADH_2$-NAD^+ charge transfer complex formed inside the catalytic site.

In our calculations, only the crystallographic water molecules with a relative accessibility to the solvent lower than 50% were retained. The relative accessibility of water was computed using the program ASC [14, 15]. The atomic partial charges and radii used in the PB calculations, for both the protein and the cofactors (FAD, $FADH_2$ and NAD^+), were derived from the GROMOS 54A7 force field [16] using the procedure described in [17]. The molecular surface was defined with a solvent probe of 1.4 Å radius and a Stern (ion-exclusion) layer of 2.0 Å. The dielectric constant was 10 for the protein/cofactors and 80 for the solvent, the temperature was 300 K and the ionic strength 0.25 M. The finite-difference linear PB calculations used a three-step focusing [18] procedure employing consecutive grid spacings of 1.0, 0.5 and 0.25 Å.

The MC calculations were done with the cofactors in fixed redox states, and with steps of 0.2 pH units. Each MC simulation comprises 10^5 MC steps and the acceptance/rejection of each step followed a Metropolis criterion [19] using the previously determined PB free energies. Each MC step consists of a first cycle of random changes of the protonation states (including tautomeric forms) of all individual sites, followed by a cycle of random double changes of the protonation states of all pairs of sites considered to be strongly coupled; a pair of sites is assumed to be strongly coupled when the

electrostatic interaction of at least one of their state combinations is above 2.0 pK_a units [10, 20].

III.2.3 Results and Discussion

In this work we searched for strictly or highly conserved structural elements of NDH-2s in order to get insights into the poorly understood catalytic mechanism of these enzymes.

NDH-2 is composed of three domains as observed in Figure III.2.1: two dinucleotide binding domains and a C-terminal domain. Coloured in green is the first dinucleotide binding domain to which the cofactor FAD is non-covalently bond. FAD's isoalloxazine ring is inside of the protein contacting the other two domains, being its re-side facing the NADH binding domain, the second dinucleotide binding domain (coloured in orange). Moreover, the first domain also includes a motif proposed to be part of the quinone binding site [6]. Each of the two dinucleotide binding domains adopts a Rossmann fold, containing the Wierenga motifs (GxGxxG patterns, or variations), known to stabilize the adenine rings of the dinucleotides [21]. The C-terminal domain (coloured in purple) contains 3-stranded antiparallel β-sheet and two amphipathic α-helices and is responsible for the association to the membrane (Figure III.2.1).

Unless otherwise mentioned, we used the amino acid sequence and tertiary structure of NDH-2 from *S. aureus* (PDB:4XDB, [5]) as reference for discussion.

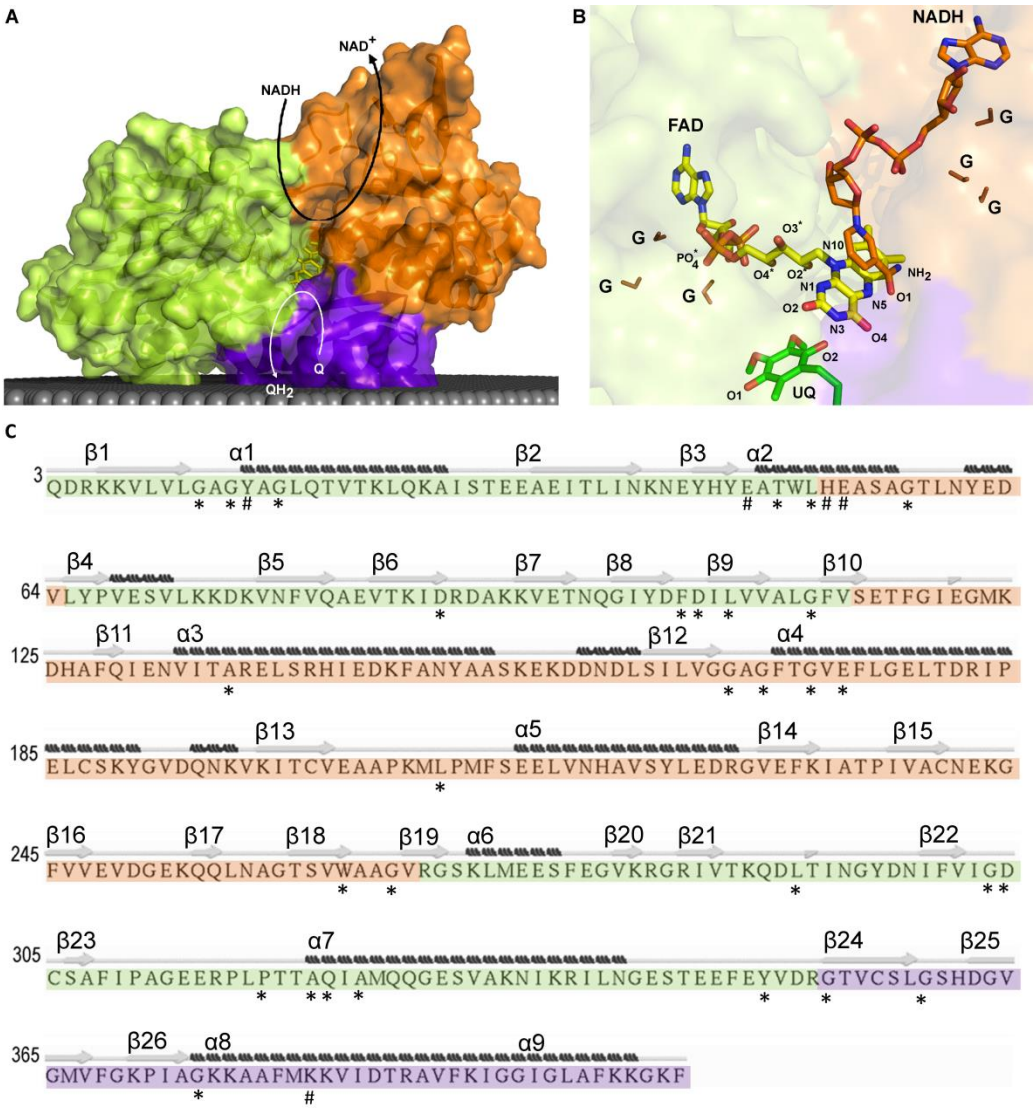


Figure III.2.1: Schematic representation of NDH-2 and substrates. NDH-2 contains three domains: first dinucleotide binding domain, or FAD binding (green); second dinucleotide binding domain or NADH binding (orange); and membrane interacting domain, amphipathic helices at the C-terminal (purple). A) Crystalized structure of NDH-2 from S. aureus. Reduction of FAD by NADH and consequent reduction of quinone by FAD; B) Oxidized FAD cofactor and co-crystalized ubiquinone and NADH of the NDH-2 from *S. cerevisiae* (PDB:4G73 [7]). The atoms of the FAD cofactor are

ordered and colored in: blue – Nitrogen atom (N); red – Oxygen atom (O); orange – Phosphorus atom (P) and yellow – Carbon atom (C). The GxGxxG motif from each dinucleotide binding domain is colored in brown; C) Secondary structure of NDH-2 from S. aureus. Secondary structure was visualized using Stride [8]. β-sheets and α-helices are numerated from the N- to the C-terminal. Residues with at least 80 % conservation (*) and with higher covariance (#) are marked.

III.2.3.1 Amino acid residues conservation in NDH-2 family

We performed a multiple sequence alignment using 1762 NDH-2s and looked for highly conserved amino acids residues. We found 30 amino acid residues with conservation equal or higher than 80 %, i.e. these amino acid residues are present in at least 80 % of the analyzed sequences.

III.2.3.1.1 First dinucleotide binding domain: FAD (and quinone) binding site(s)

Eighteen conserved amino acid residues are present in the 1st dinucleotide binding domain (Figures III.2.1 and III.2.2). From the N-terminal to the C-terminal of the NDH-2 sequence the first conserved motif observed is GxGxxG ($G_{12}xG_{14}xxG_{17}$ in *S. aureus*). In 9 % of NDH-2s, the last glycine residue of this motif is substituted by an alanine residue, as in the case of the NDH-2 from *S. cerevisiae* (Ndi1) [7, 22]. This glycine rich motif is in a loop located close to the pyrophosphate moiety of the FAD and should stabilize it, since permits close contact of the protein main chain to the pyrophosphate moiety (Figure III.2.1) [6, 21, 23].

A

	S. aureus NDH-2 sequence		Alignment of 1762 sequences from NDH-2 family		
	Amino acid number	Amino acid type	Amino acid Conservation (%)	Amino acid Consensus	Proposed function
1st Rossmann domain	12	G	100	G	FAD stabilization
	14	G	100	G	FAD stabilization
	17	G	83	G	FAD stabilization
	48	T	87	P	Structural / H-bond with Y_{15}
	50	L	91	L	-
	87	D	88	D	-
	102	F	92	Y	Structural / Regulation
	103	D	96	D	Structural / Regulation
	105	L	98	L	Structural
	110	G	100	G	FAD stabilization
	289	L	82	L	-
	301	G	100	G	H-bond with Q_{325}
	302	D	100	D	H-bond with O3* (FAD)
	316	P	83	P	H-bond with N1 and O2 (FAD)
	319	A	96	A	Quinone binding site motif
	320	Q	84	Q	Quinone binding site motif
	322	A	98	A	Quinone binding site motif
	347	Y	82	Y	H-bond with Q_{320}

B

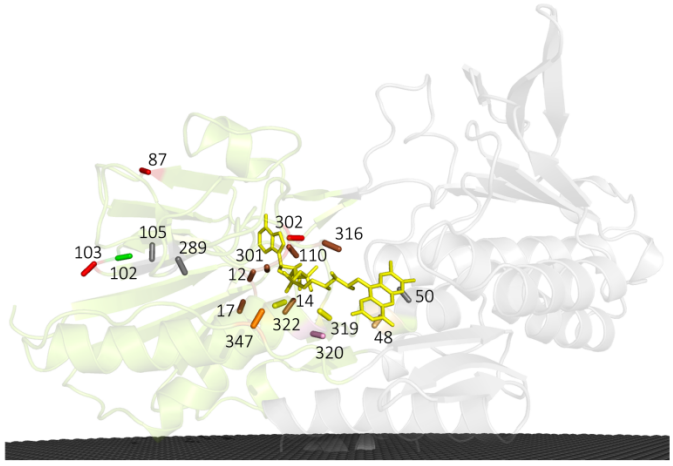


Figure III.2.2: Amino acid conservation in the 1st dinucleotide binding domain. A) 18 amino acids that are conserved in at least 80 % of the NDH-2s; B) Localization of the 18 conserved amino acids in the 1st dinucleotide (FAD) binding domain. Membrane is presented in black. NDH-2 from *S. aureus* (PDB code: 4XDB, [5]) was used for the schematic representation.

The second conserved motif, localized at the surface of the protein (Figure III.2.2), is composed of the residues pair, YD ($F_{102}D_{103}$ in *S. aureus*, Y and D are present in 92 % and 96 % of the NDH-2s, respectively). We observed that the tyrosine residue correspond to a phenylalanine residue in 7 % of the NDH-2s, as in the case of NDH-2 from *S. aureus* (F_{102}). The conservation of the YD pair is intriguing since it is localized far away from the active centre and binding sites of the two substrates. As this pair is present between 2-stranded β-sheets (Figure III.2.1, C, β8 and β9), it may be hypothesised that the hydrophilic character of the D_{103} that points to the solvent may constrain the localization of the β9 (part of the Rossmann fold) and therefore this pair may have a structural function for the Rossmann fold. Alternatively, the conserved YD pair may constitute a signal site for regulation of the enzymatic activity.

Two other conserved residues are L_{105} and G_{110}. These residues seem to have a structural role (Figure III.2.2), L_{105} may be important for the stability of the Rossmman fold by its aliphatic and G_{110} may have a role in the stabilization of pyrophosphate moiety of FAD (Figure III.2.1 and III.2.2).

Close to O3* from FAD (Figure III.2.1, B), a third strictly conserved pair, $G_{301}D_{302}$, was observed (Figure III.2.2). G_{301} in *S. aureus* establishes a backbone hydrogen bond (2.7 Å) with the side chain of a glutamine residue. This glutamine (Q_{325}) is localized after the proposed quinone binding site motif ($A_{319}Q_{320}xA_{322}xQ_{324}$ in *S. aureus*) [6] in α-helix 7 (Figure III.2.1, C). We noticed the presence of such a bond in all three available structures of NDH-2s [5-7]. This residue corresponds to

a glutamine in 57 % of the NDH-2s and corresponds to a glutamate and to histidine residues in the cases of NDH-2 from *S. cerevisiae* and *C. thermarum*, respectively. D_{302} in *S. aureus* was previously suggested to make hydrogen bonds with FAD, through both its backbone to the PO_4 group and its side chain to O3* from FAD (Figure III.2.1, B) [6, 22, 24, 25]. Studies performed with NDH-2 from *S. cerevisiae*, in which the equivalent aspartate residue was mutated (by an alanine, an asparagine, a glutamine or a glutamate) had shown that the presence of glutamate/aspartate residue is important for the activity of the enzyme [26]. In addition the same aspartate residue was proposed to be part of the quinone binding site [26]. The high conservation of that aspartate residue (D_{302}) is extended to several families of the tDBDF superfamily, even in those whose members do not interact with quinones (our results, unpublished data), suggesting it could be important in the processes of oxidation / reduction of the FAD (and not because it is part of the quinone binding site).

We further observed that 3 out of the 4 first residues from the quinone binding site motif, AQxAxQ [1, 6], have a conservation higher than 80 % (Figure III.2.2), while the last glutamine residue is still present in 78 % of NDH-2s. The backbone of A319 and Q320 were described before as making direct hydrogen bond interactions with the isoalloxazine ring of FAD [6, 22], while the two glutamine residues (Q_{320} and Q_{324} in *S. aureus*) were proposed to be at the entrance to the active site being responsible for the delimitation of hydrophobic and hydrophilic sites [6, 22]. We also investigated alternative quinone binding site motifs and observed the presence of three main motifs:

AQxAxQ (already mentioned above), AQxAxR (wherein the last glutamine residue is replaced by an arginine residue) and APxAxQ (wherein the first glutamine residue is replaced by a proline residue). These three motifs are conserved in 62 %, 15 % and 10 % of the 1762 NDH-2s, respectively [1].

III.2.3.1.2 Second dinucleotide binding domain: NADH binding site

The second dinucleotide binding domain harbours the NADH binding site (Figure III.2.1). In this domain, we identified nine conserved residues forming different motifs (Figures III.2.1 and III.2.3). The first conserved motif is localized at the beginning of α-helix 4 and is GxGxxGxE (G_{165}xG_{167}xxG_{170}xE_{172} in *S. aureus*). The glycine residues were proposed to stabilize the adenine moiety of NADH [6, 7, 26] and the glutamate residue is at hydrogen/electrostatic bond distance from the co-crystalized NADH nicotinamide ring in the structure of NDH-2 from *S. cerevisiae* [7] (Figures III.2.1 and III.2.3).

Replacing the first glycine residue by a serine residue hampered *S. cerevisiae* to grow [26]. The glutamate residue (E_{172} in *S. aureus*, E_{169} in *C. thermarum* and E_{242} in *S. cerevisiae*) is conserved in 97 % of the 1762 NDH-2s, while a glutamine residue is present in the remaining sequences (3 %, mainly Archaea). Single mutation experiments showed that yeast cells had growth defects when E_{242} (E_{172} in *S. aureus*) was replaced by an alanine or aspartate residues [7]. This mutation also affected NADH and quinone kinetic parameters (K_M and V_{max}) [7], suggesting an important role of this residue in the catalytic mechanism of NDH-2.

A

	S. aureus NDH-2 sequence		Alignment of 1762 sequences from NDH-2 family		
	Amino acid number	Amino acid type	Amino acid Conservation (%)	Amino acid Consensus	Proposed function
2nd Rossmann domain	56	G	81	G	-
	134	A	90	A	-
	165	G	100	G	NAD(P)H stabilization
	167	G	99	G	NAD(P)H stabilization
	170	G	99	G	NAD(P)H stabilization
	172	E	97	E	NAD(P)H stabilization / Part of a proton pathway (2nd RD)
	207	L	92	L	-
	261	W	99	W	NAD(P)H stabilization
	264	G	100	G	NAD(P)H stabilization

B

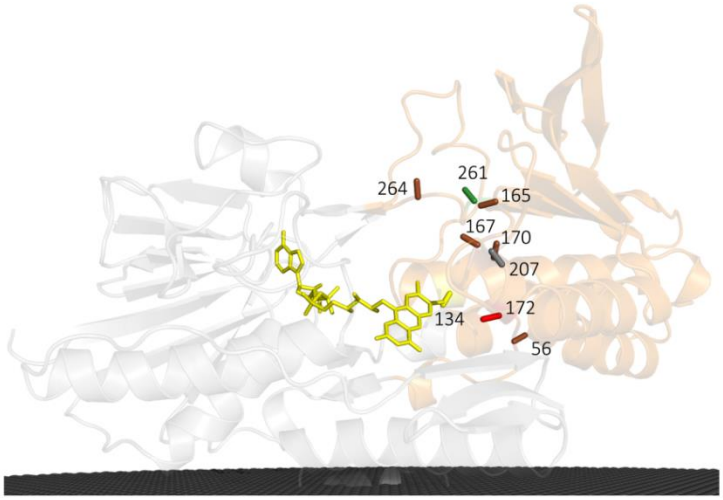


Figure III.2.3: Amino acid conservation in the 2nd dinucleotide binding domain. A) 9 amino acids are conserved in at least 80 % of the NDH-2s; B) Localization of the 9 conserved amino acids in the 2nd dinucleotide binding domain. Membrane is presented in black. NDH-2 from *S. aureus* (PDB code: 4XDB, [5]) was used for the schematic representation.

The motif WxxG ($W_{261}xxG_{264}$ in *S. aureus*) is also highly conserved, 99 % and 100 %, respectively (Figures III.2.1 and III.2.3). As W_{261} is close to NADH adenine base we hypothesize its importance in the orientation and/or stabilization of NAD(P)H.

III.2.3.1.3 C-terminal domain: Membrane interacting module

The C-terminal domain allows protein interaction with the membrane through two amphipathic α-helices 8 and 9, as observed in the structures of NDH-2s from three different organisms (Figure III.2.1) [5-7, 22]. We found three conserved glycine residues (G_{351}, G_{357} and G_{372}, Figure III.2.4). The last glycine residue (G_{372}) is conserved in 99 % of NDH-2s and we hypothesize its presence is important to orient the first amphipathic α-helix to the catalytic centre. By comparing the crystallographic structures of the two quinone reducing families (NDH-2 [5-7, 22] and SQR [27, 28]) of the tDBDF superfamily we observed that, in both cases, the first amphipathic α-helix occupies the same position, close to the isoalloxazine ring of the FAD. The structural localization of this α-helix allows the side chains of its amino acid residues to interact with FAD and respective substrates (NAD(P)H or sulfide and quinone).

III.2.3.2 Covariance analysis of amino acid residues in NDH-2 family

Aiming not to exclude possible relevant amino acid residues with lower conservation (not retrieved in our first analysis), we performed a covariance analysis using the MISTIC tool [4]. This tool gives insights into the relation between two residues by predicting positional correlations based on the structure and multiple sequence alignment. For example during evolution, an amino acid residue at position “A” important for the redox activity can be changed without loss of the protein activity if a change in another amino acid residues at position “B” takes place, compensating for the change of the first amino acid at position “A”.

A

	S. aureus NDH-2 sequence		Alignment of 1762 sequences from NDH-2 family		
	Amino acid number	Amino acid type	Amino acid		Proposed function
			Conservation (%)	Consensus	
C-terminal domain	351	G	99	G	Structural
	357	G	80	G	Structural
	372	G	99	G	Structural / Orientation of α9

B

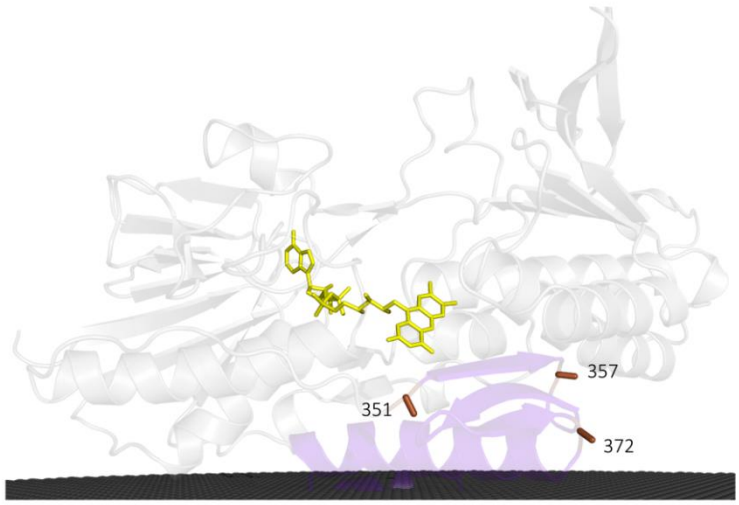


Figure III.2.4: Amino acid conservation in the C-terminal domain. A) 3 amino acids are conserved in at least 80 % of the NDH-2s; B) Localization of the 3 conserved amino acids in the C-terminal domain. Membrane is presented in black. NDH-2 from *S. aureus* (PDB:4XDB, [5]) was used for the schematic representation.

Our analysis revealed the existence of residues with high cumulative covariance, i.e. sum of all relations between a residue at a certain position to the others residues positions. We selected all residues with cumulative covariance above 70 %, when normalizing in relation to the positions with higher cumulative covariance, which were X_{15} and X_{379} (100% of cumulative covariance). Therefore, we accepted for analysis three residues positions besides X_{15} and X_{379}

(Figure III.2.5). Note that the residue positions are referred to the nomenclature of the NDH-2 from *S. aureus*.

III.2.3.2.1 Covariance in the first dinucleotide binding domain (FAD binding site)

The first dinucleotide binding domain showed two positions (X_{15} and X_{46}) included in the 5 positions with the highest cumulative covariance (Figures III.2.1 and III.2.5). X_{15} (Y_{15} in *S. aureus*) is part of the FAD binding motif, $G_{12}xG_{14}Y_{15}xG_{17}$ (Figure III.2.2). This position is occupied by an aromatic residue in 81 % of NDH-2s, varying between a phenylalanine (35 %), a tyrosine (18 %, Y_{13} in C*. thermarum*) or a tryptophan (28 %, W_{63} in *S. cerevisiae*) residue. In 16 % of the cases, the conserved aromatic character is lost and replaced by an alanine residue (Figure III.2.5). X_{15} was previous described as being part of the tunnel extending from the C-terminal domain to the *si*-side of the FAD and able to establish a direct hydrogen bond, through its backbone, with one of the PO_4 groups from FAD (Figure III.2.1, B) [6, 22]. In the NDH-2´scrystallographic structures [5-7, 22] is observed that the side chain of X_{15} establishes a hydrogen bond with a threonine residue (T_{48} in *S. aureus*, T_{46} in *C. thermarum* and T_{91} in *S. cerevisiae*). This analysis also showed that X_{15} is highly correlated with X_{51}, X_{355}

A

	S. aureus NDH-2 sequence				
	Residue number	Residue type	Residue consensus	Covariance (%)	Proposed function
1st Rossmann domain	15	Y	F	100	FAD stabilization / H-bond with T_{48}
	46	E	F	76	Part of a proton pathway (1st RD)
2nd Rossmann domain	51	H	Y	73	Part of a proton pathway (2nd RD)
	52	E	E	74	Part of a proton pathway (2nd RD)
C-terminal domain	379	K	W	100	Part of a proton pathway (2nd RD) / H-bond with O4 (FAD) / protonation of the quinone

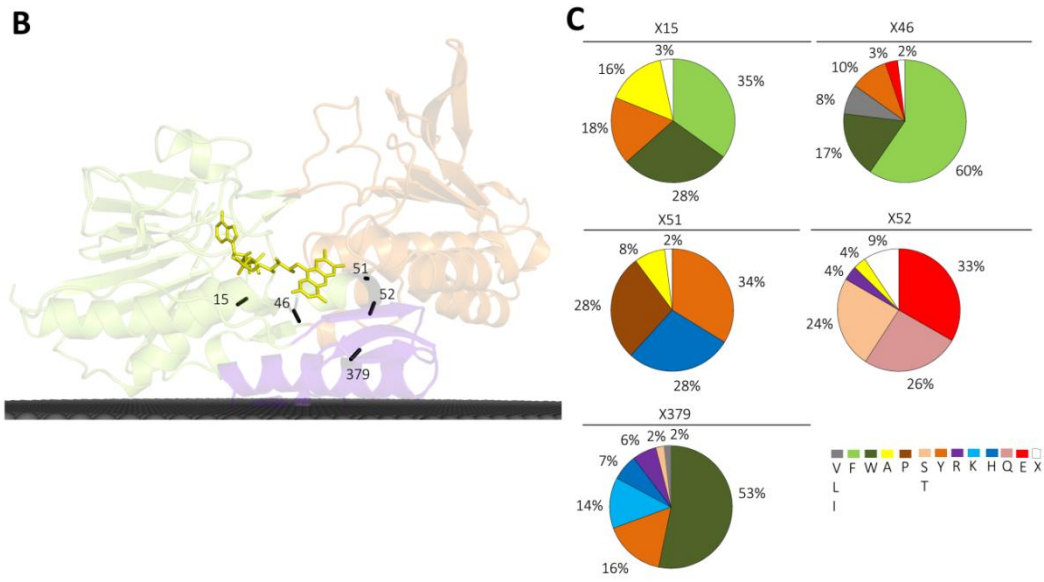


Figure III.2.5: Amino acid residues covariance in NDH-2 family. A) 5 amino acid residues positions with higher cumulative covariance (above 70 %); B) Localization of the 5 amino acid residues positions with higher cumulative covariance. C) Amino acid residues frequency at the 5 amino acid positions with higher covariance. Membrane is presented in black. NDH-2 from *S. aureus* (PDB:4XDB, [5]) was used for the schematic representation.

and X_{379} (corresponding to H_{51}, S_{355} and K_{379} in *S. aureus*). Interestingly, these three positions are at hydrogen/electrostatic bond distance from the highly conserved E_{172} (see below). X_{46} (E_{46} in *S. aureus*, see below), another position with a high cumulative covariance, is located at the *si*-side of FAD (Figures III.2.1 and III.2.5). This position is

occupied by an aromatic residue (phenylalanine, tyrosine or tryptophan) in 87 % of the NDH-2s (Figure III.2.5) and also showed high covariance with X_{49}, X_{168} and X_{379} (corresponding to W_{49}, F_{168} and K_{379} in *S. aureus*).

III.2.3.2.2 Covariance in the second dinucleotide binding domain (NADH binding domain)

In the second dinucleotide binding domain, two residues localized at the *re*-side of FAD (H_{51} and E_{52} in *S. aureus*), are included in the 5 positions with the highest cumulative covariance (Figures III.2.1 and III.2.5). As can be seen in Figure III.2.5, X_{51} (H_{51} in *S. aureus*) vary mainly within three residues: tyrosine (34 %), histidine (28 %, H_{49} in *C. thermarum*) or proline (28 %, P_{95} in *S. cerevisiae*), while X_{52} (E_{52} in *S. aureus*) may contain a glutamate (33 %), glutamine (26 %, Q_{50} in *C. thermarum*) or serine (24 %, S_{96} in *S. cerevisiae*) residues. We observed that a glutamate residue (two residues before) is also present in the vicinity of the histidine (H_{49}) in the case of *C. thermarum* (corresponding to E_{47} in *C. thermarum*). These residues pairs (H_{51} and E_{52} in *S. aureus* and E_{47} and H_{49} in *C. thermarum*) seem to form a conserved motif that may have a role in the proton transfer process (the two residues composing the pair are at ~3.3 Å and ~3.9 Å apart, respectively). In NDH-2s from *S. aureus* and *C. thermarum*, H_{51} is also at hydrogen bond distance from the side chain of the highly conserved E_{172}, from the side chain of K_{379} (at ~3.3 Å) and near N5 from the FAD isoalloxazine ring (~6.8 Å in *S. aureus* and see Figure III.2.1, B). The analysis of protonation equilibrium simulations performed for *S.*

aureus evidenced that H_{51} is sensitive to the redox state of the FAD cofactor (Figure III.2.6, A).

We also observed that X_{51} and X_{52} have high covariance with X_{15} and X_{355}, and to X_{15} and X_{383}, respectively (corresponding to Y_{15}, S_{355} and D_{383} in *S. aureus*).

III.2.3.2.3 Covariance in the C-terminal domain (membrane interacting module)

X_{379} (K_{379} in *S.* aureus) is localized at the C-terminal domain and is the last of the 5 positions with the highest cumulative covariance (Figures III.2.1 and III.2.5, purple domain). X_{379} is a tryptophan residue in 53 % (W_{478} in *S. cerevisiae*), a positively charged residue (K, H or R) in 27 % (K_{376} in *C. thermarum*), or a hydroxyl containing residue (16 % tyrosine and 2 % threonine) (Figure III.2.5). We also observed that X_{379} has higher covariance with X_{15}, X_{46} and X_{168} (Y_{15}, E_{46} and F_{168} in *S. aureus*).

III.2.3.3 Identification of two distinct proton pathways in NDH-2

All the catalytic steps in NADH:quinone oxidoreduction, i.e. NADH oxidation, FAD reduction and oxidation and quinone reduction involve proton transfers and in this way we looked for possible proton pathways by further exploring the conservation of amino acids by type (e.g. protonable and aromatic,) and analysis of the three available NDH-2 structures [5-7, 22]. We were able to identify two distinct proton pathways.

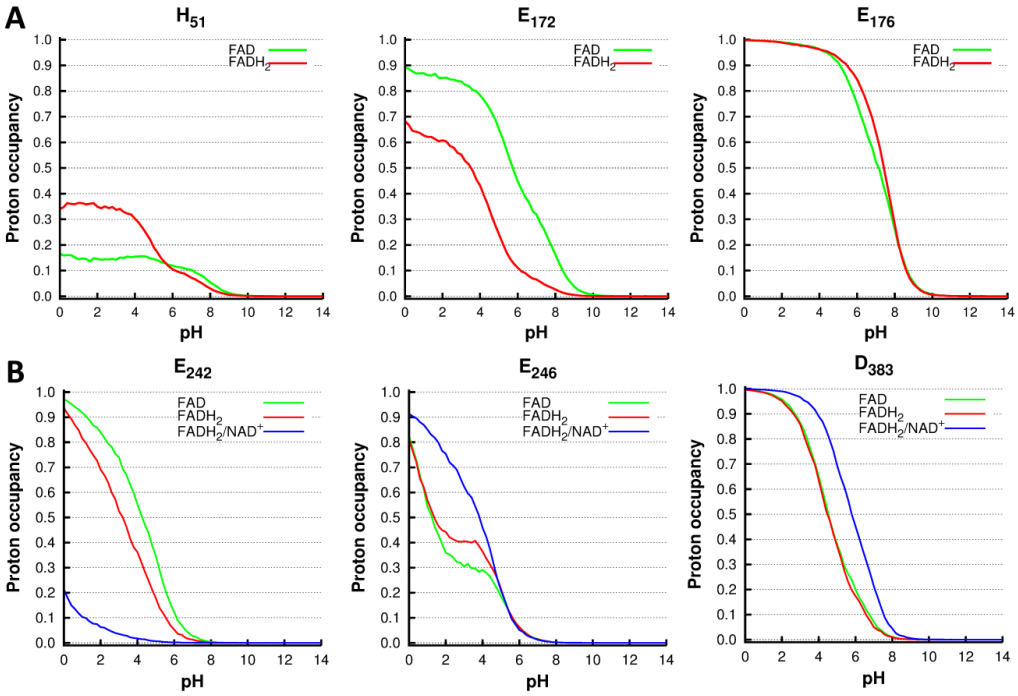


Figure III.2.6: Simulated pH titrations for the NDH-2 from *S. aureus* (A) and *S. cerevisiae* (B). In these plots, we show the pH titration curves for all the residues whose protonation behaviour shows dependence with the FAD redox state.

III.2.3.3.1 A proton pathway in the second dinucleotide binding domain (NADH binding site)

On the *re*-side of the FAD, we observed that the conserved E_{172} (E_{169} in *C. thermarum* and E_{242} in *S. cerevisiae*) is at hydrogen/electrostatic bond distance from several residues and from the -NH_2 group of the nicotinamide ring of NAD(P)H when it is bond to the protein (Figure III.2.1, B). We noticed that this glutamate is localized on the interior side of a wire composed of carboxylate residues connected to the surface of the protein in the three NDH-2

structures (Figure III.2.7). All these carboxylate residues are oriented to the same side of α-helix 4 (Figure III.2.1, C). These residues are E_{172}, E_{176}, D_{179} and E_{183} in *S. aureus* (Figure III.2.7, A), E_{169}, E_{173}, D_{176} and E_{180} in *C. thermarum* (Figure III.2.7, B) and E_{242}, E_{246}, D_{249} and D_{254} in *S. cerevisiae* (Figure III.2.7, C), and have an overall conservation of 97 %, 62 %, 74 % and 28 %, considering the conservation for the carboxylate residues (glutamate or aspartate).

We further identified the shorter connections between the surface and the conserved $E_{172}/E_{169}/E_{242}$ for the three structures of NDH-2 (Figure III.2.7). We observed the following combinations: in *S. aureus*, E_{183} to D_{179} (6.2 Å, at the surface), D_{179} to backbone of S_{358} (3.3 Å), backbone of S_{358} to E_{176} (4.1 Å), E_{176} to S_{355} (3.5 Å), S_{355} to H_{51} (2.2 Å) and H_{51} to E_{172} (3.7 Å); in *C. thermarum*, E_{180} to D_{176} (9.6 Å, at the surface), D_{176} to E_{173} (5.4 Å), E_{173} to R_{355} (1.8 Å), R_{355} to H_{49} (3.1 Å) and H_{49} to E_{169} (3.8 Å) and in *S. cerevisiae*, D_{254} to R_{507} (2.0 Å, at the surface), R_{507} to D_{249} (3.8 Å) and D_{249} to Y_{449} (2.4 Å). This Y_{449} is at 5.2 Å from E_{246} and at 9.4 Å from E_{242}, since in the NDH-2 from *S. cerevisiae* (and most of eukaryotic NDH-2s) no histidine is present to connect the last two glutamates (E_{242} and E_{246}) from the proton conducting wire. This connection could be hypothesized to be performed under a rotamer change state of Y_{449} (corresponding to S_{355} in *S. aureus*); or conformational changes at S_{96} (corresponding to E_{52} in *S. aureus*) since there are two proline residues before that could unwound the α-helix 2 (Figure III.2.1) and move S_{96} closer to E_{242}; or by the presence of a

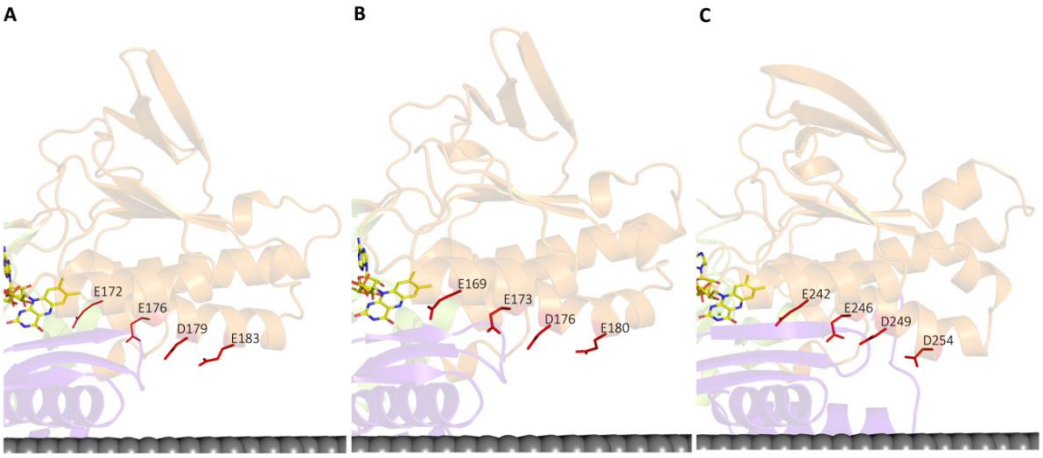


Figure III.2.7: Proton pathways in the second dinucleotide binding domain. A) Wire of carboxylate residues in the NDH-2 from *S. aureus* (PDB:4XDB, [5]); B) Wire of carboxylate residues in the NDH-2 from *C. thermarum* (PDB:4NWZ [6]) and C) Wire of carboxylate residues in the NDH-2 from *S. cerevisiae* (PDB:4G73 [7]). Membrane is presented in black.

structural water molecule, which curiously is observed between $P_{95,}$ E_{242} and E_{246} residues in both NDH-2 structures of *S. cerevisiae* [7, 22].

The protonation equilibrium simulations performed for *S. aureus* clearly shows that the protonation of E_{172} is highly dependent on the redox state of the FAD. As can be seen in Figure III.2.6, panel A, when the FAD molecule is oxidized, E_{172} shows a protonated fraction of 0.32 at pH 7. However, when the cofactor is reduced, the E_{172} protonated fraction decreases significantly to 0.06. Thus, E_{172} is likely to play a role in proton transfer during the catalytic cycle. In the *S. cerevisiae* enzyme, a similar dependence between the protonation of E_{242} (residue equivalent to $E_{172\ in}$ *S. aureus)* and the redox state of the FAD cofactor is also observed (Figure III.2.6, B).

In our calculations, E_{176} has a significant protonated population at pH 7 (0.52 and 0.64 when cofactor is oxidized and reduced, respectively), being also slightly affected by the redox state of the FAD molecule (Figure III.2.6, A). Similarly, to what was described for the *S. aureus* enzyme, E_{246} in *S. cerevisiae* is also affected by the reduction of the FAD (Figure III.2.6, B).

In summary, a wire of residues can be hypothesised as a proton pathway from the protein surface to the NADH pocket (near E_{172}/E_{169}/E_{242}), which we assume that has an important role in proton transfer during the catalytic cycle (see below).

III.2.3.3.2 A proton pathway in the first dinucleotide binding domain (FAD binding site)

In the structure of NDH-2 from *S. cerevisiae* we observed a possible proton conductive wire composed of residues present in α-helices 1 and 7 (Figure III.2.1, C), which could supply protons to the quinone pocket. This proton pathway may involve: E_{401} at 3.5 Å from H_{397}, H_{71} at 4.4 Å from E_{401}, and three more residues that could interact with H_{71} upon rearrangement of the side chain (D_{73}, K_{405} and D_{408}) (Figure III.2.8, C).

Interestingly, $O1_q$ atom from the co-crystalized quinone is at 5.4 Å from H_{397} (Figures III.2.1, B and III.2.8, C). This histidine residue is present at the quinone binding site motif, AQxAH_{397}Q, in the NDH-2

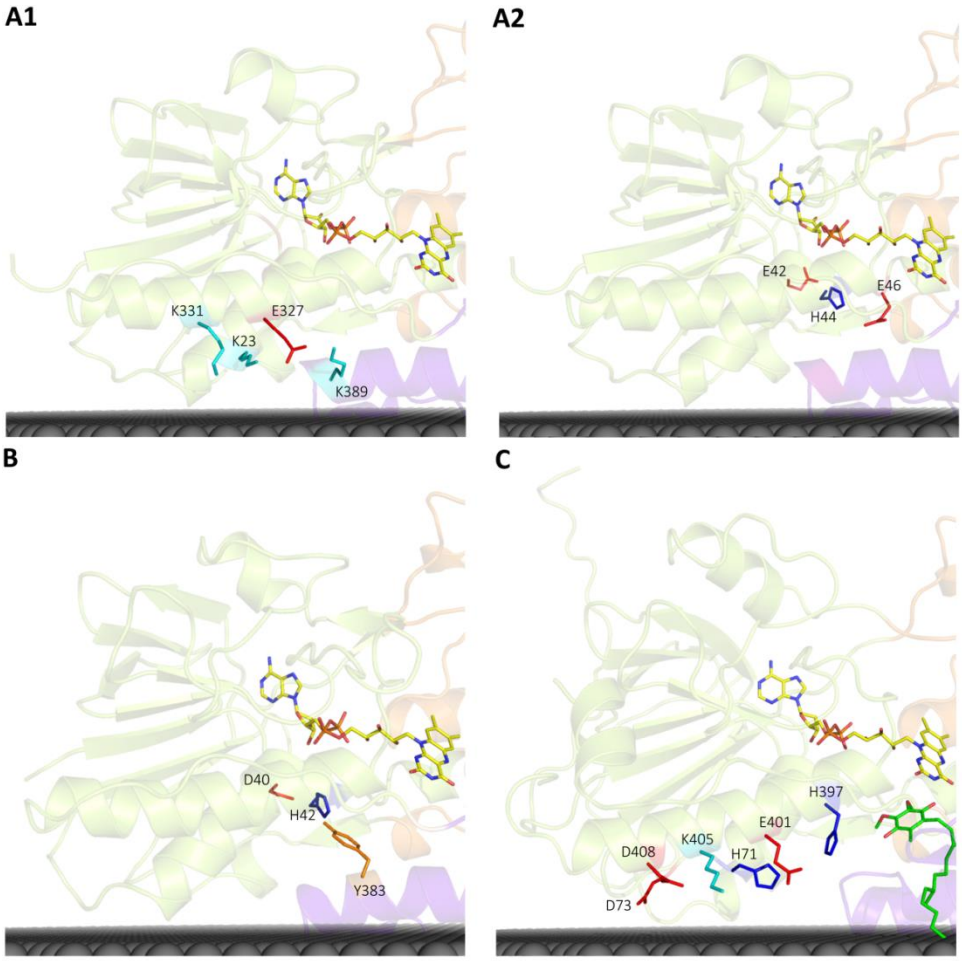


Figure III.2.8: Proton pathways in the first dinucleotide binding domain. A1) proton pathway trough α-helices 1 and 7 of the NDH-2 from *S. aureus* (PDB:4XDB, [5]), A2) Alternative proton pathway in the NDH-2 from *S. aureus* (PDB:4XDB, [5]), B) Only proton pathway found for NDH-2 from *C. thermarum* (PDB:4NWZ [6]), C) proton pathway trough α-helices 1 and 7 of the NDH-2 from *S. cerevisiae* (PDB:4G73 [7]). Membrane is presented in black.

from *S. cerevisiae*. H_{397} of NDH-2 from *S. cerevisiae* is only present in 17 % of NDH-2s, mainly in proteobacteria and some eukaryotic species. We observed 6 %, 8 % and 9 % of the NDH-2s contain a lysine,

a glutamate or a serine residue, respectively, that may act as proton conductor at equivalent position of H_{397} in *S. cerevisiae*.

In the NDH-2s from *S. aureus* and *C. thermarum*, a methionine ($AQxAM_{323}Q$) and an isoleucine ($AQxAI_{320}Q$) residue, respectively, are observed at the position equivalent to H_{397} of NDH-2 from *S. cerevisiae*. This may suggest that the quinone binding pocket for menaquinone reducing NDH-2s is different or another residue could replace the histidine present in *S. cerevisiae.* Considering the same quinone binding pocket, we searched for other residues in the vicinity that could be relevant. We noticed two positions whose amino acid residues side chain may spatially occupy the side chain position of H_{397} in NDH-2 from *S. cerevisiae*. The first position is X_{19} (*S. aureus* nomenculature), after the first conserved $G_{12}xGxxG_{17}$ motif (Figure III.2.2), where X_{19} can be occupied by a lysine (3 %), an arginine (9 %), a tyrosine (14 %) or a glutamate (18 %). The second position is in the C-terminal of the 1st amphipathic α-helix (Figure III.2.1, C, α8), where a lysine residue (X_{389} corresponding to K_{389} in *S. aureus*) and a tyrosine (one loop before at X_{386}, corresponding to Y_{383} in C*. thermarum*) are present in the NDH-2s from *S. aureus* and *C. thermarum*, respectively (Figure III.2.8, A1 and B). We observed that 7 % of the NDH-2s have a lysine, an arginine, a histidine or a tyrosine residue at X_{389}, while 58 % NDH-2s have these residues at X_{386}. We explored the presence of a proton conductive wire from the K_{389} to the protein surface in NDH-2 from *S. aureus* and we noticed a wire composed of residues also localized in α-helices 1 and 7 (Figure III.2.1). This proton wire is composed of: E_{327} at 3.0 Å from K_{389}, K_{23} at 2.0 Å from E_{327}, K_{331} at 2.9 Å

from K_{23}. K_{331} is localized at the surface of NDH-2 from *S. aureus* (Figure III.2.8, A1). Alternatively, three residues may form also a proton wire to the quinone binding pocket in NDH-2 from *S. aureus*. This alternative pathway is composed of H_{44} (present in 53 % of the NDH-2s) and E_{46} (Figure III.2.8, A2), which are at 2.3 Å from each other.

In the case of *C. thermarum*, we did not find a proton pathway trough α-helices 1 and 7 as in the case of *S. cerevisiae*, but we observed also the presence of a histidine (H_{42}) in the same place of H_{44} from NDH-2 from *S. aureus* (Figure III.2.8, A2 and B). This time, we observe an isoleucine (I_{44}) two positions after, where a glutamate (E_{46}) is present in *S. aureus* (Figure III.2.8, A2). However, the tyrosine (Y_{383}) present at the C-terminal of NDH-2 from *C. thermarum* makes a hydrogen bond with H_{42} (2.9 Å), which suggests that the tyrosine may play the same role of the E_{46} from *S. aureus* (Figure III.2.8, A2 and B). Moreover, we observed the presence of a glutamate or an aspartate residue (at the surface of NDH-2) two positions before the histidine (H_{44}/H_{42}) in both NDH-2s (E_{42}/D_{40}) (Figure III.2.8, A2 and B). These residues are at more than 10 Å from the histidine (H_{44}/H_{42}) but could have a role in connecting the histidine to the surface upon conformational changes since it is located in a loop or by the presence of water molecules. This alternative proton pathway seems to not exist in *S. cerevisiae* since no histidine is present and a tryptophan present in the $GxGxW_{64}G$ motif seems to block that location to the quinone binding pocket.

In contrast to the proton wire observed in the 2nd dinucleotide binding domain that is conserved among NDH-2 family and is

composed mainly of glutamate/aspartate residues (Figure III.2.7), the proton conductive wire present in the 1st dinucleotide binding domain seems to adopt different residues arrangements (Figure III.2.8). We hypothesize that this difference on the proton conductive wire of the 1st dinucleotide binding domain may be related to the type of quinone, menaquinone as in the case of *S. aureus* and *C. thermarum* or ubiquinone as the case of *S. cerevisiae*.

The highly conserved glycine-rich Wierenga motifs (GxGxxG) [21] present in both dinucleotide binding domains (see above section III.2.3.1.1 and III.2.3.1.2) are localized in the beginning of α-helices 1 and 4, which are (partially) part of the proposed proton pathways (Figures III.2.7 and III.2.8). This suggest that conformational changes upon reduction/oxidation of FAD and binding/oxidation of NAD(P)H, which could be reflected at the Wierenga motifs, may induce proton transfer through the two proposed proton pathways.

III.2.3.4 Hypothesis for the catalytic mechanism of NDH-2

Despite the available crystallographic data and kinetic studies [5-7, 22, 26, 29], the catalytic mechanism of NDH-2 is still unclear.

Oxidation of NADH involves the release of two electrons and deprotonation of the its nicotinamide ring (Figure III.2.9, A). The two electrons are transferred to FAD, in a process usually accompanied by FAD protonation at positions N1 and N5 of its isoalloxazine ring (Figures III.2.1, B and III.2.9, B). The subsequent reduction of quinone comprises the transfer of two electrons from FAD to quinone, deprotonation of FAD and protonation of the two reactive oxygen

atoms present in the quinone ($O1_q$ and $O2_q$, Figures III.2.1, B and III.2.9, C). The mechanism is believed to involve the establishment of a ternary complex. This hypothesis is supported by the observation in NDH-2 from *S. aureus* using visible fast kinetic analyses, of a charge-transfer complex formed between NAD^+ and the reduced flavin (FADH2), which was dissociated by the quinone [5]. In the case of the enzyme from *E. coli*, the establishment of a ternary is sustained by the observation that both substrates were bond to the enzyme at the same time [29].

Here, we discuss the possible roles of the conserved elements in the catalytic process, including proton transfer and substrate interaction. We divided the discussion in two parts: FAD reduction (by NADH) and FAD oxidation (by quinone).

III.2.3.4.1 FAD reduction (first half reaction)

The way by which FAD is reduced in NDH-2 is unknown, but based on what is observed for several flavoproteins we propose that FAD is reduced by hydride transfer from NADH (at the re-side) [30]. In this way, N5 from FAD isoalloxazine can accept the hydride from C4 of the nicotinamide ring of NADH (which is at ~3.4 Å in NDH-2 structure from *S. cerevisiae)* (Figure III.2.9).

The origin of the second proton needed for the full protonation of FAD is not clear. Three hydrogen-bonding interactions between FAD and highly conserved amino acid residues called our attention. Two of these interactions, are present between N1 and O2 atoms of the oxidized isoalloxazine ring (Figures III.2.1, B and III.2.9, D) and the

backbone of the two first residues from the quinone binding site ($A_{319}Q_{320}$xAxQ in *S. aureus*) and the third hydrogen-bonding interaction is observed between O3* from FAD (Figure III.2.1, B) and the side chain of the conserved aspartate residue, part of the GD motif described previously (D_{302} in *S. aureus* – see above section III.2.3.1.1) (Figure III.2.9, D).

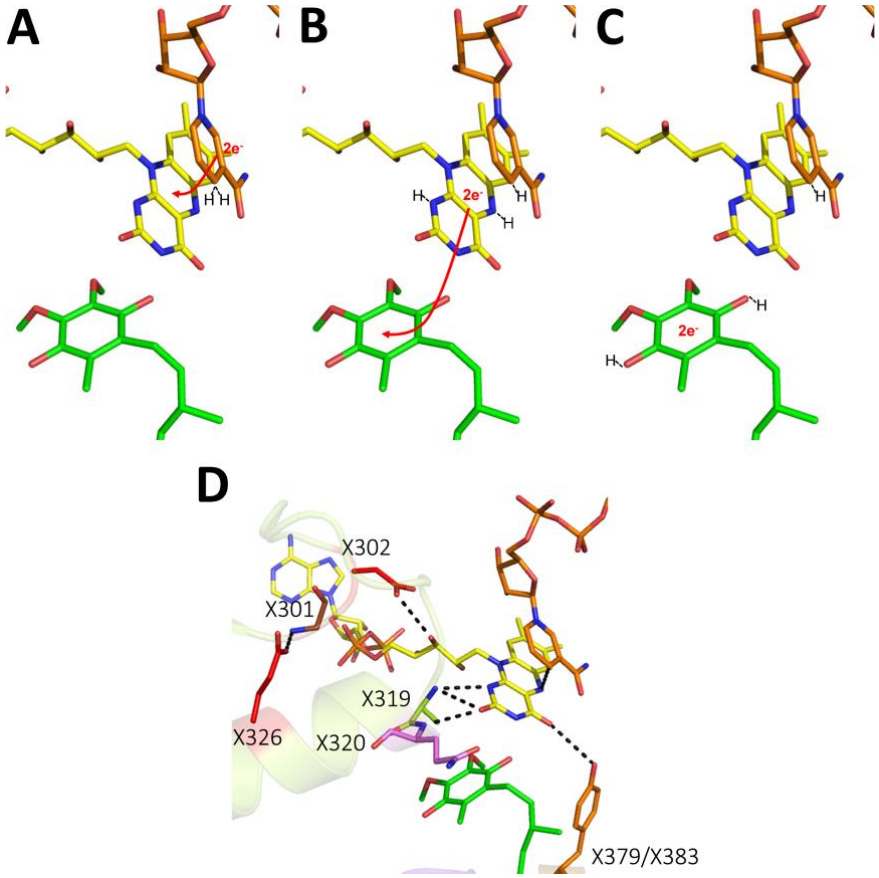


Figure III.2.9: NADH:quinone oxidoreductase activity involves protons and electrons transfer. A) NADH (reduced), FAD (oxidized) and quinone (oxidized). NADH transfers 2 electrons to FAD and releases one proton. The reduction of FAD can be

performed by hydride transfer from NADH. B) NAD^+ (oxidized), $FADH_2$ (reduced) and quinone (oxidized). $FADH_2$ transfers 2 electrons to quinone and releases two protons. The reduction of quinone can be performed by hydride transfer from $FADH_2$. C) NADH (oxidized), FAD (oxidized) and quinone (reduced). Reduced quinone and oxidized NAD^+ are released from NDH-2. D) Some of the residues involved in the stabilization of FAD, where dashed lines are hydrogen bonds: backbone of X_{301} (G residue) interacts with X_{326} (E/Q/H in *S. cerevisiae*, *S. aureus*, *C. thermarum*, respectively); X_{302} (D) interacts with O3* from FAD; backbones of X_{319} and X_{320} interact with N1 and O2 from FAD isoalloxazine ring ; X_{379}/X_{383} (X_{379} is a K in *S. aureus* and *C. thermarum* and X_{383} is a Y in *S. cerevisiae*) interact with O4 from FAD isoalloxazine ring. Schemes based on the NDH-2 co-crystalized with NADH and quinone from *S. cerevisiae* (PDB:4G73 [7]).

The protonation equilibrium simulations performed in the *S. cerevisiae* enzyme showed that the protonation equilibrium of D_{383} (equivalent to D_{302} in *S. aureus*) is greatly influenced by the presence/absence of NAD^+ in the catalytic site (Figure III.2.6, B). When the NAD^+ is absent, the protonated fraction of D_{383} is almost negligible at pH 7 (0.07 and 0.05 for the FAD and the $FADH_2$ states). However, when NAD^+ is forming a complex with $FADH_2$, the protonated fraction of D_{383} substantially increases to 0.21 (Figure III.2.6, B). This may indicate that this residue is involved in proton network influencing the protonation of FAD at N1 site.

Since the D_{302}/D_{383} is at ~7-8 Å of N1 site, a direct protonation event is not possible, which suggests the involvement of other elements on the proton transfer process between the aspartate residue and the N1 site. As the members of the tDBDF superfamily are structurally similar and likely share the same proton interaction

between FAD and conserved residues. We investigated the structures of two of its members in the reduced state, NADH:ferredoxin oxidoreductase and reduced thioredoxin reductase, (also members of the tDBDF superfamily) in order to identify this extra the possible element involved in proton conduction. In the referred structures, the isoalloxazine ring of the reduced FAD adopts a bent conformation (the so-called boat conformation), which contrasts with the planar conformation observed in the oxidised FAD. The best conformation with causes the rotation of O2* (Figure III.2.1, B), which promotes the connection between N1 and an aspartate (corresponding to D_{302} in *S. aureus*) residue allowing the proton transfer between them [24, 25].

In summary, FAD may be reduced by NADH through hydride transfer and its second protonation may be achieved by rearrangement of the hydrogen bond network due to a change in FAD conformation upon reduction.

III.2.3.4.2 FAD oxidation (second half reaction)

FAD is oxidised by the quinone (interacting at the *si*-side) which has to be protonated at each side of its reactive oxygen atoms ($O1_q$ and $O2_q$, Figures III.2.1, B and III.2.9, C). Resembling the reduction of FAD reduction by hydride transfer from NADH, the simplest way for quinone reduction is also by hydride transfer from N5 of the isoalloxazine ring of FAD, rather than N1 of FAD to the quinone because the mechanism of deprotonation of N1 is conserved in the members of the tDBDF superfamily, which includes non-interacting quinone enzymes.

In the structure of NDH-2 from *S. cerevisisae*, N5 is at 6.3 Å from the nearest reactive oxygen ($O2_q$) of the quinone (Figure III.2.1, B). Considering that $O2_q$ is reduced/protonated by hydride transfer from N5 of FAD (which will be further discussed in section IV.4), still one proton is needed to protonate $O1_q$ position of the quinone (and another proton has to be released from N1 of FAD). We propose that when FAD is reoxidised by the quinone, the FAD readopts its conformation of the oxidised state, losing the bent conformation and consequently the hydrogen bond network involving the GD motif and O2* of FAD, which results in deprotonation of N1.

For the protonation of $O1_q$ from the quinone, we suggest that the nearest residue belonging to the proton pathways present in the 1st dinucleotide binding domain is the best candidate to provide protons to the quinone. H_{397} may play this role in the case of NDH-2 from *S. cerevisiae*, because it is at the inner end of the proton pathway that we proposed to be present in the 1st dinucleotide binding domain (Figure III.2.8, C). In the case of NDH-2 from *S. aureus* we observed two possible proton pathways (Figure III.2.8), one with the same arrangement of that of NDH-2 from *S. cerevisiae* (Figure III.2.8, A and C), where the last residue (K_{389} in *S.aureus*) is near the quinone binding site motif and an alternative proton pathway that includes H_{44} and E_{46} Figure III.2.8, B). Both sets of residues (K_{389} or H_{44}/E_{46}) may protonate $O1_q$. In relation to NDH-2 from *C. thermarum* no glutamate residue (E_{46} in *S. aureus*, Figure III.2.8, A2) is present in that position, instead a tyrosine (Y_{383} in *C. thermarum,* Figure III.2.8, B) or a water molecule may have the same role. We propose that this tyrosine under rotamer

change or a water molecule (which is present in the crystalized structure of *C. thermarum*) (Figure III.2.8, B) may protonates the $O1_q$ of quinone.

In summary, we propose that FAD may be oxidised by the quinone through hydride transfer (see an alternative hypothesis in section IV.4) and its second deprotonation may be achieved by rearrangement of the hydrogen bond network upon a conformational change of FAD losing it boat conformation induced by reoxidation. The quinone may get the second proton through the proton conductive pathway present in the 1st dinucleotide binding domain.

III.2.3.4.3 What is the role of the highly conserved E_{172}?

The mechanism just described above is sufficient to explain the oxidation and reduction processes as well as the protonation events. However, we noticed the presence of E_{172} in 97 % of NDH-2s, which is located close to the NADH binding site and seems to be at the end of a proton conductive pathway connected to the protein surface. Furthermore, our calculations indicate the protonation state of E_{172} (E_{242} in *S. cerevisiae*) is highly dependent on the oxidation state of FAD (Figure III.2.1, C and III.2.6). Thus the intriguing question now is what is the role of the highly conserved E_{172} in the catalytic mechanism?

The side chain of E_{172} (E_{169} in *C. thermarum* and E_{242} in *S. cerevisiae*) may establish three possible different hydrogen bonds with residues in its vicinity, namely with H_{51} (which protonation state is dependent on the oxidation state of FAD, Figure III.2.6), backbone of S_{355} or K_{379} in the NDH-2 from *S. aureus* (Figure III.2.10, A1); with H_{49},

backbone of S_{352} or K_{376} in the NDH-2 from *C. thermarum* (Figure III.2.10, B1) and with backbone of Y_{449}, W_{478} or Y_{482} in the NDH-2 from *S. cerevisiae* (Figure III.2.10, C1). These residues are included in the proposed proton pathways for the 2nd dinucleotide binding domain (see section IV.3.2) or are residues with higher covariance (X_{51}/X_{379}/X_{383} in section III.2.3.2).

We propose the side chain of the glutamate (E_{172}/E_{169}/E_{242}) may establish an interaction with the NH_2 group of NADH (at ~4.4 Å in *S. cerevisiae*) upon NADH/NAD^+ binding (Figure III.2.10, A). This interaction may influence the other interactions of E_{172}/E_{169}/E_{242} with the residues in the vicinity (Figure III.2.10), allowing the other residues to establish new interactions.

In summary, upon the formation of the charge transfer complex $FADH_2$-NAD^+, a hydrogen/electrostatic bonds between NAD^+ and the conserved glutamate (E_{172}/E_{169}/E_{242}) may be established, disrupting hydrogen/electrostatic bonds of this glutamate with the residues in the vicinity. This proposal suggests that the conserved glutamate plays a role as a gate in the hydrogen/electrostatic bonds networks from residues belonging to the proton pathway of the 2nd dinucleotide binding domain (see below).

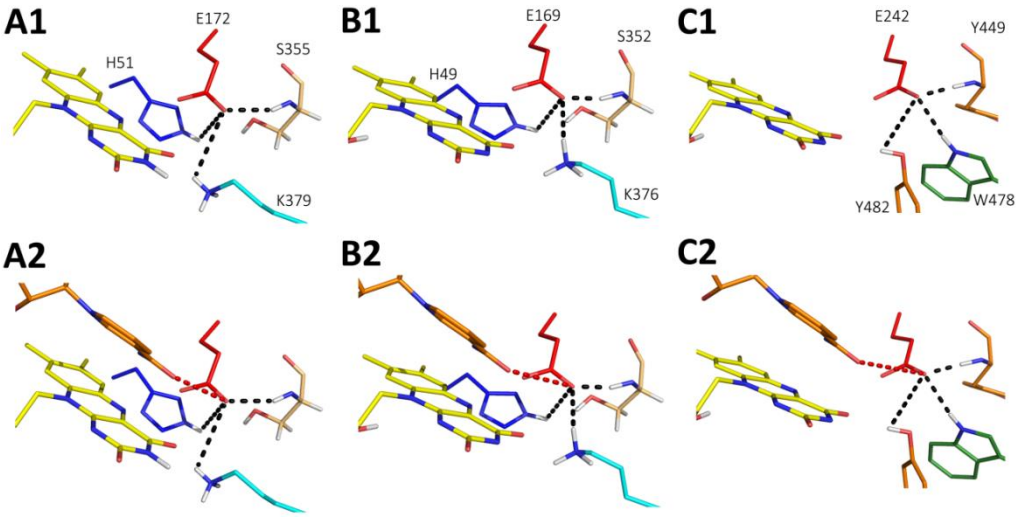


Figure III.2.10: Role of glutamate (E_{172} in *S. aureus*) in the catalytic mechanism. Dash lines show possible hydrogen/electrostatic bonds with the glutamate. A1) E_{172} in *S. aureus*. E_{172} may interact with H_{51}, S_{355} or K_{379} when FAD is oxidized (black dashed lines). B1) E_{169} in *C. thermarum*. E_{169} may establish interact with H_{49}, S_{352} or K_{376} when FAD is oxidized (black dashed lines). C1) E_{242} in *S. cerevisiae*. E_{242} may interact with Y_{449}, W_{478} or Y_{482} when FAD is oxidized (black dashed lines). A2, B2, and C2) Upon binding of NADH and $FADH_2$-NAD^+ complex formation is established possibly a new hydrogen/electrostatic bond with NH_2 from NADH (red dash lines) and others disrupted (black dash line). Schemes based on the NDH-2 co-crystalized with NADH and quinone from *S. cerevisiae* (PDB:4G73 [7]) and NDH-2s from *S. aureus* (PDB:4XDB, [5]) and *C. thermarum* (PDB:4NWZ, [6]).

III.2.3.4.4 Residues X_{51} and X_{379}/X_{383}: Are these protons conducting elements?

E_{172} may interact with K_{379} through a hydrogen bond (~3.3 Å in *S. aureus*), which is the position with the highest cumulative covariance (X_{379}, Figure III.2.5). K_{379} is located close to the isoalloxazine ring of FAD (at ~3.3 Å from O4 from FAD) and at the interface between

the quinone and NADH pockets. We observed the presence of a proton conducting residue at X_{379} in 45 % of the NDH-2s, including *S. aureus* (K_{379}) and *C. thermarum* (K_{376}), while in 53 % of the NDH-2s a tryptophan is present (W_{478} in *S. cerevisiae*) (Figures III.2.5 and III.2.11).

NDH-2 from *S. cerevisiae* is one of the 53 % proteins that contains a tryptophan residue (W_{478}) at the equivalent position to X_{379}. Interestingly, we observed in its structure the presence of a tyrosine residue (Y_{482}) four positions ahead (next loop), whose side chain seems to occupy the same structural position as that of K_{379} from *S. aureus* (structural alignment between NDH-2s from *S. aureus* and *S. cerevisiae* [RMSD = 1.2 Å]). In the structure of NDH-2 from *S. cerevisiae*, Y_{482} is at ~2.5 Å from O4 of FAD isoalloxazine ring and at ~3.1 Å from the highly conserved E_{242} (Figure III.2.10, C).

We further noticed for the 53 % proteins that contain a tryptophan residue at the equivalent position to X_{379}, 51 % and 46 % have a tyrosine or a histidine residue at X_{383} (corresponding to Y_{482} in *S. cerevisiae*), respectively (Figure III.2.11). Thus when analyzing the X_{379} and X_{379+4} (X_{383}) positions in the NDH-2 alignment, we observed 98 % NDH-2s have a proton conducting residue at the interface of the NADH and quinone pockets (X_{379}/X_{383}, Figure III.2.11).

Moreover, in the case of *S. aureus*, E_{172} (X_{172}) may also interact with H_{51} (X_{51}), which are at ~4.2 Å apart, which is also included in the residues with high cumulative covariance. X_{51} varies in 90 % of the NDHs-2 between a tyrosine, a histidine or a proline (Figure III.2.5 and III.2.11).

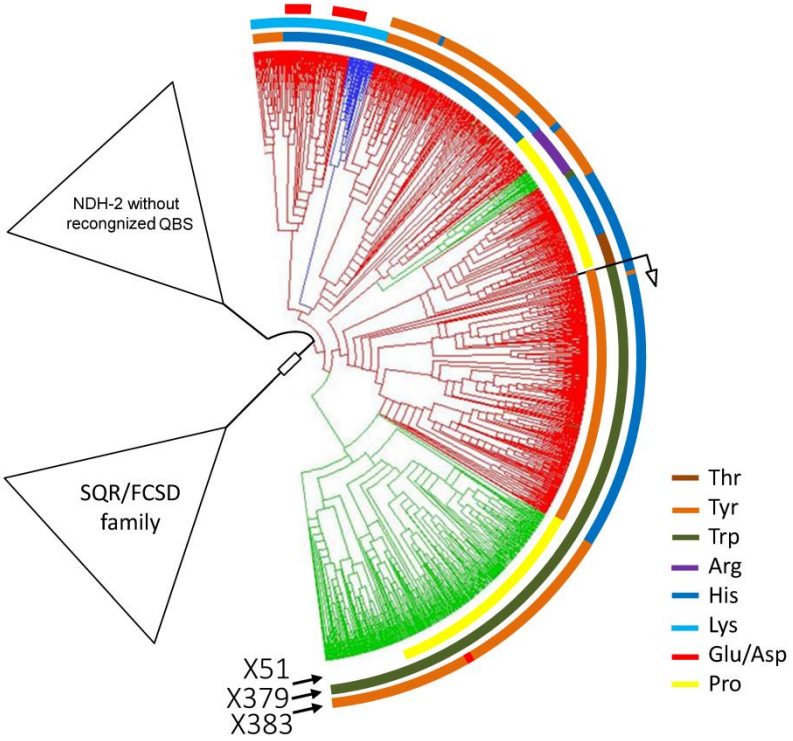


Figure III.2.11: Neighbor-joining dendrogram of NDH-2s family. The dendrogram was obtained using 2567 sequences from NDH-2s and sequences from SQRs/FCCDs to define the outer group, midpoint root, 1000 replicates of bootstrap, 999 generated seeds [1]. Branches are colored according to the three domains of life from which the organism containing the respective NDH-2 originates: green corresponds to microorganisms from Eukarya, red to Bacteria and blue to Archaea. The external ring of the dendrogram, in the region of NDH-2s, indicates the distribution of three residues proposed to be involved in proton transfer (X_{51}, X_{379} and X_{383}). The type of residue is indicated in the legend of the figure.

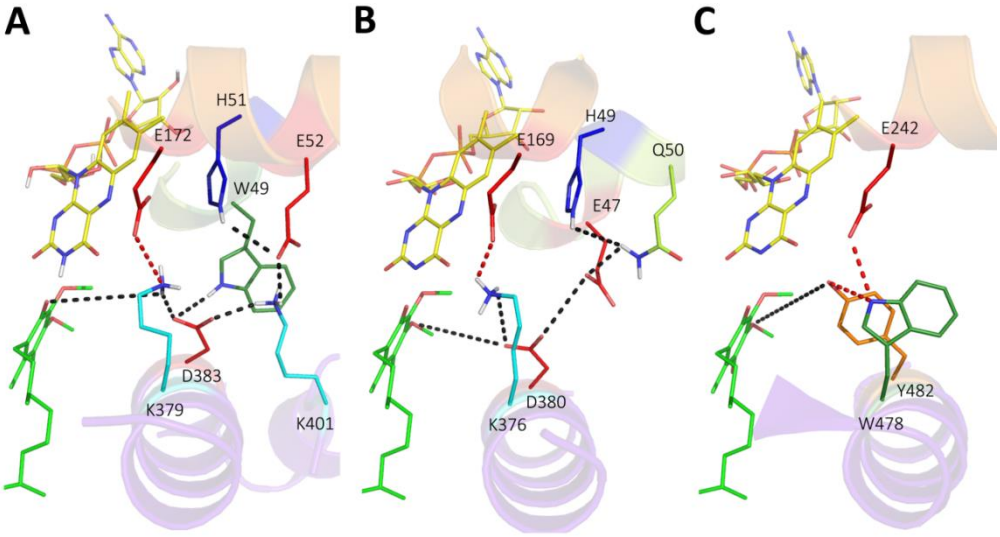


Figure III.2.12: Proton conducting elements from X_{172} (E_{172}) to the quinone binding pocket. Red dash lines show direct proton transfer to X_{379}/ X_{383}. Black dash lines show alternative proton transfer to X_{379}/ X_{383} trough X_{51}. A) In the NDH-2 from *S. aureus* E_{172} could release a proton to K_{379} or a proton could be transferred trough H_{51}- E_{52}- K_{401}- D_{383}. B) In the NDH-2 from *C. thermarum* E_{169} could release a proton to K_{376} or a proton could be transferred trough H_{49}- Q_{50}- E_{47}- D_{380}. C) In the NDH-2 from *S. cerevisiae* E_{242} could release a proton to Y_{482} trough W_{478}. Schemes based on the NDH-2 co-crystalized with NADH and quinone from *S. cerevisiae* (PDB:4G73 [7]) and NDH-2s from *S. aureus* (PDB:4XDB, [5]) and *C. thermarum* (PDB:4NWZ, [6]).

We observed that H_{51} (in *S. aureus*) can be connected to the quinone binding pocket, through E_{52} (~3.3 Å), E_{52} to K_{401} (~2.8 Å) and K_{401} to D_{383} (~2.4 Å). The latter seems to be stabilized by K_{379} (~2.8 Å) and W_{49} (~3.5 Å) (Figure III.2.12, A).

In the case of *C. thermarum*, H_{49} (X_{51}) can be linked via Q_{50} (~3.9 Å), Q_{50} to E_{47} (~2.5 Å) and E_{47} to D_{380} (~3.8 Å). D_{380} may be stabilized by K_{376} at ~3.7 Å (Figure III.2.12, B). In the case of *S. cerevisiae* a proline is

present at X_{51}, therefore we observed only one possible pathway through a hydrogen bond between E_{242} and Y_{482} (Figure III.2.12, C).

In order to assign the described network as a pathway for the protonation of $O2_q$ of quinone (in alternative to the described hydride transfer between FAD and quinone – see section III.2.3.4.2), we extended our analyses to SQRs, which are the closest members of the tDBDF superfamily to the NDH-2 family and the only ones to share quinone as substrate. We performed a structural alignment (RMSD=3.2 Å) using the SQR crystallographic structure from *Aquifex aeolicus* (PDB:3HYW, [27]) and that of NDH-2 from *S. aureus* [5]. We observed SQR from *A. aeolicus* also have the proposed connections from position X_{172} (E_{172} in *S. aureus*) to the quinone binding pocket, the following residues being superimposable: E_{172} with E_{162}, H_{51} with H_{47}, side chain of K_{401} with R_{53}, K_{379} with K_{382} and D_{383} with E_{386}. We also observed that E_{52} is not present in SQR from *A. aeolicus* but at X_{355} (S_{355} in *S. aureus*) there is an aspartate (D_{350}) that seems to play the same role. Moreover, the lysine residue at X_{379} was suggested to be a proton donor to the quinone [27, 28]. We observe that the lysine residue (K_{382} in *A. aeolicus*) is at least 62 % conserved among SQRs [1].

These results reinforce the proposal for the presence of a proton conducting residue at the interface of NADH/sulfide and quinone pockets for these two families. As both families share the same electron acceptor we may hypothesize that the residues referred here (X_{51} and X_{379}/X_{383}) may have a role in quinone protonation, possibly as a proton conducting elements.

III.2.4 Conclusions

We performed an exhaustive bioinformatic analysis in order to identify the conserved amino acid residues and structural elements within NDH-2 family. We carried out this analysis with NDH-2s with recognized quinone binding motifs, *i.e.* ~70 % of the 2567 proteins considered members of the NDH-2 family [1]. We identified 30 amino acid residues conserved in at least 80 % of the NDH-2 sequences and we recognized five positions with high cumulative covariance (X_{15}, X_{46}, X_{51}, X_{52} and X_{379}). Combining the conservation analysis and the structural information of the three available structures from NDH-2s [5-7, 22], we were able to identify relevant elements, such as a proton pathway in each of the dinucleotide binding domains. The proton pathway from the 2nd dinucleotide binding domain (NADH binding) is more conserved among the NDH-2 family than that observed in the 1st nucleotide binding domain and is composed by several glutamates or aspartates residues. The proton pathway present in the 1st dinucleotide binding domain (FAD binding) seems to adopt different geometries, which may reflect it is specialization to different types of quinone, menaquinone or ubiquinone. Both proton pathways connect protons from the surface of the protein to the quinone pocket. The localization of the two proton pathways suggests that the quinone pocket can be feed by protons at both sides of the quinone reactive oxygens (considering the co-crystalized quinone in the NDH-2 structure from *S. cerevisiae* [7]). Moreover, the highly conserved E_{172} (present in 97 % of NDH-2 sequences) seems to be part of the proton pathway present at the 2nd dinucleotide binding domain (NADH binding) and may have a role in the coordination of the proton

transfer. We suggest that E_{172} by interacting with NH_2 group from nicotinamide ring of NADH may alter hydrogen/electrostatic bonds with amino acid residues present in the vicinity, namely X_{51} (when is a histidine or a tyrosine) and X_{379} or X_{383} (when X_{379} is a tryptophan residue). The change in hydrogen/electrostatic bonds may induce conformational changes allowing proton transfer to $X_{51}/X_{379}/X_{383}$.

As observed for other members from the tDBDF, we suggest that upon reduction of FAD by NADH, FAD undergoes conformational change that affects conserved residues at the 1st dinucleotide binding domain (FAD binding), namely the conserved GD and the quinone binding site motifs. The adjustment of residues for the stabilization of the reduced FAD, may induce alterations in the β3 and α-helices 1 and 7. Interestingly, the proposed quinone binding site (α7) and proposed proton pathway (β3 or α1 and α7) are present in the referred α-helices.

Curiously, amino acid sequence insertions (EF-hand or CxxC motifs, see [1]) between the conserved residues that form the GD motif and the next α-helix (α7) are observed in several NDH-2s [1]. The EF-hand motif was suggested to bind calcium and to regulate the NDH-2 activity [31]. These motifs may constitute sites for regulation of FAD cofactor by inducing local conformational changes in the residues that stabilize/protonate FAD in a certain oxidation state.

Finally, the distribution of NDH-2s based on key residues such as X_{51}, X_{379} and X_{383} may be relevant for understanding the type of quinone used in the catalytic reaction of NDH-2 and may give insights

of the pathways where NDH-2 is involved since several species code for more than one type of NDH-2.

III.2.5 Acknowledgments

We acknowledge M. Teixeira, R. O. Louro and P. Refojo for the critical reading of the manuscript. FVS is recipient of a fellowship by Fundação para a Ciência e a Tecnologia (PD/BD/113985/2015). The work was funded by Fundação para a Ciência e a Tecnologia (PTDC/BBB-BQB/2294/2012 to M.M.P). ITQB is supported by Fundação para a Ciência e a Tecnologia through R&D Unit, UID/CBQ/04612/2013.

III.2.6 References

1. Marreiros, B. C., Sena, F. V., Sousa, F. M., Batista, A. P. & Pereira, M. M. (2016) Type II NADH:Quinone oxidoreductase family: Phylogenetic distribution, Structural diversity and Evolutionary divergences, *Environmental microbiology*.

2. Pei, J., Kim, B. H. & Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments, *Nucleic acids research.* 36, 2295-300.

3. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench, *Bioinformatics.* 25, 1189-91.

4. Simonetti, F. L., Teppa, E., Chernomoretz, A., Nielsen, M. & Marino Buslje, C. (2013) MISTIC: Mutual information server to infer coevolution, *Nucleic acids research.* 41, W8-14.

5. Sena, F. V., Batista, A. P., Catarino, T., Brito, J. A., Archer, M., Viertler, M., Madl, T., Cabrita, E. J. & Pereira, M. M. (2015) Type-II NADH:quinone oxidoreductase from

Staphylococcus aureus has two distinct binding sites and is rate limited by quinone reduction, *Molecular microbiology.* 98, 272-288.

6. Heikal, A., Nakatani, Y., Dunn, E., Weimar, M. R., Day, C. L., Baker, E. N., Lott, J. S., Sazanov, L. A. & Cook, G. M. (2014) Structure of the bacterial type II NADH dehydrogenase: a monotopic membrane protein with an essential role in energy generation, *Molecular microbiology.* 91, 950-64.

7. Feng, Y., Li, W., Li, J., Wang, J., Ge, J., Xu, D., Liu, Y., Wu, K., Zeng, Q., Wu, J. W., Tian, C., Zhou, B. & Yang, M. (2012) Structural insight into the type-II mitochondrial NADH dehydrogenases, *Nature.* 491, 478-82.

8. Heinig, M. & Frishman, D. (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins, *Nucleic acids research.* 32, W500-2.

9. Teixeira, V. H., Soares, C. M. & Baptista, A. M. (2002) Studies of the reduction and protonation behavior of tetraheme cytochromes using atomic detail, *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry.* 7, 200-16.

10. A.M. Baptista, C. M. S. (2001) Some theoretical and computational aspects of the inclusion of proton isomerism in the protonation equilibrium of proteins, *J Phys Chem B.* 105, 293-309.

11. Bashford, D. & Gerwert, K. (1992) Electrostatic calculations of the pKa values of ionizable groups in bacteriorhodopsin, *Journal of molecular biology.* 224, 473-86.

12. Bashford, D. & Karplus, M. (1990) pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model, *Biochemistry.* 29, 10219-25.

13. D, B. (1997) *An Object-Oriented Programming Suite for Electrostatic Effects in Biological Molecules*, Springer, Berlin: ISCOPE97.

14. Polticelli, F., Bottaro, G., Battistoni, A., Carri, M. T., Djinovic-Carugo, K., Bolognesi, M., O'Neill, P., Rotilio, G. & Desideri, A. (1995) Modulation of the catalytic rate of Cu,Zn superoxide dismutase in single and double mutants of conserved positively and negatively charged residues, *Biochemistry.* 34, 6043-9.

15. F. Eisenhaber, P. A. (1993) Improved strategy in analytic surface calculation for molecular-systems - handling of singularities and computational efficiency, *J Comput Chem.* 14 . 1272-1280.

16. Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E. & van Gunsteren, W. F. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7, *European biophysics journal : EBJ.* 40, 843-56.

17. Teixeira, V. H., Cunha, C. A., Machuqueiro, M., Oliveira, A. S., Victor, B. L., Soares, C. M. & Baptista, A. M. (2005) On the use of different dielectric constants for computing individual and pairwise terms in poisson-boltzmann studies of protein ionization equilibrium, *The journal of physical chemistry B.* 109, 14691-706.

18. M.K. Gilson, K. S., B. Honig, R. Fine, R. Hagstrom (1987) Calculation of electrostatic energies in proteins by a finite-difference method, *Biophysical journal.* 51, A234-A234.

19. N. Metropolis, A. R., M. Rosenbluth, A. Teller (1953) Equation of state calculations by fast computing machines, *J Chem Phys.* 21, 1087-1092.

20. Baptista, A. M., Martel, P. J. & Soares, C. M. (1999) Simulation of electron-proton coupling with a Monte Carlo method: application to cytochrome c3 using continuum electrostatics, *Biophysical journal.* 76, 2978-98.

21. Wierenga, R. K., Terpstra, P. & Hol, W. G. (1986) Prediction of the occurrence of the ADP-binding beta alpha beta-fold in proteins, using an amino acid sequence fingerprint, *Journal of molecular biology.* 187, 101-7.

22. Iwata, M., Lee, Y., Yamashita, T., Yagi, T., Iwata, S., Cameron, A. D. & Maher, M. J. (2012) The structure of the yeast NADH dehydrogenase (Ndi1) reveals overlapping binding sites for water- and lipid-soluble substrates, *Proceedings of the National Academy of Sciences of the United States of America.* 109, 15247-52.

23. Wierenga, R. K., Drenth, J. & Schulz, G. E. (1983) Comparison of the three-dimensional protein and nucleotide structure of the FAD-binding domain of p-hydroxybenzoate hydroxylase with the FAD- as well as NADPH-binding domains of glutathione reductase, *Journal of molecular biology.* 167, 725-39.

24. Lennon, B. W., Williams, C. H., Jr. & Ludwig, M. L. (1999) Crystal structure of reduced thioredoxin reductase from Escherichia coli: structural flexibility in the

isoalloxazine ring of the flavin adenine dinucleotide cofactor, *Protein science : a publication of the Protein Society.* 8, 2366-79.

25. Senda, M., Kishigami, S., Kimura, S., Fukuda, M., Ishida, T. & Senda, T. (2007) Molecular mechanism of the redox-dependent interaction between NADH-dependent ferredoxin reductase and Rieske-type [2Fe-2S] ferredoxin, *Journal of molecular biology.* 373, 382-400.

26. Yang, Y., Yamashita, T., Nakamaru-Ogiso, E., Hashimoto, T., Murai, M., Igarashi, J., Miyoshi, H., Mori, N., Matsuno-Yagi, A., Yagi, T. & Kosaka, H. (2011) Reaction mechanism of single subunit NADH-ubiquinone oxidoreductase (Ndi1) from Saccharomyces cerevisiae: evidence for a ternary complex mechanism, *The Journal of biological chemistry.* 286, 9287-97.

27. Marcia, M., Ermler, U., Peng, G. & Michel, H. (2009) The structure of Aquifex aeolicus sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration, *Proceedings of the National Academy of Sciences of the United States of America.* 106, 9625-30.

28. Cherney, M. M., Zhang, Y., Solomonson, M., Weiner, J. H. & James, M. N. (2010) Crystal structure of sulfide:quinone oxidoreductase from Acidithiobacillus ferrooxidans: insights into sulfidotrophic respiration and detoxification, *Journal of molecular biology.* 398, 292-305.

29. Salewski, J., Batista, A. P., Sena, F. V., Millo, D., Zebger, I., Pereira, M. M. & Hildebrandt, P. (2016) Substrate-Protein Interactions of Type II NADH:Quinone Oxidoreductase from Escherichia coli, *Biochemistry*.

30. Ojha, S., Meng, E. C. & Babbitt, P. C. (2007) Evolution of function in the "two dinucleotide binding domains" flavoproteins, *PLoS computational biology.* 3, e121.

31. Melo, A. M., Duarte, M., Moller, I. M., Prokisch, H., Dolan, P. L., Pinto, L., Nelson, M. A. & Videira, A. (2001) The external calcium-dependent NADPH dehydrogenase from Neurospora crassa mitochondria, *The Journal of biological chemistry.* 276, 3947-51.

Chapter IV

Exploring membrane respiratory chains

Chapter IV – Exploring membrane respiratory chains

IV.1 – Exploring membrane respiratory chains

This section is based on the following publication:

Marreiros BC, Calisto F, Castro PJ, Duarte AM, Sena FV, Silva AF, Sousa FM, Teixeira M, Refojo PN and Pereira MM (2016), "Exploring membrane respiratory chains" Biochimica et Biophysica Acta Bioenergetics.

DOI: 10.1016/j.bbabio.2016.03.028.

Note that the results presented in this section are summarized in order to not further enhance the PhD thesis' number of pages. Figure IV.1.1 summarizes the taxonomic profile of all the respiratory membrane bound complexes presented in this chapter.

Author contribution:

B.C.M. participated in the design of the study, performed and analyzed the experiments, critically discussed the data and drafted the manuscript.

IV.1 – Exploring membrane respiratory chains

IV.1.1 Summary

Acquisition of energy is central to life. Organisms need energy for the establishment and maintenance of a transmembrane difference in electrochemical potential allowing cells to import and export metabolites or for their motility. The membrane potential is established by a variety of membrane-bound respiratory complexes. In this work we explored the diversity of membrane respiratory chains and the presence of the different enzyme complexes in the several phyla of life. We performed taxonomic profiles of the several membrane-bound respiratory proteins and complexes evaluating the presence of their respective coding genes in all species deposited in KEGG database. We evaluated 26 quinone reductases, 5 quinol:electron carriers oxidoreductases and 18 terminal electron acceptor reductases. We further included in the analyses enzymes performing redox or decarboxylation driven ion translocation, ATP synthase and transhydrogenase and we also investigated the electron carriers that perform functional connection between the membrane complexes, quinones or soluble proteins. Our results bring a novel, broad and integrated perspective of membrane-bound respiratory complexes and thus of the several energetic metabolisms of living systems.

IV.1.2 Material and Methods

IV.1.2.1 Protein databases and taxonomic profiling

Genes coding for proteins or protein complexes belonging to the different families involved in membrane respiratory chains were searched among Eukarya, Bacteria and Archaea domains using protein BLAST (BLASTp) analysis tool running at KEGG (Kyoto Encyclopedia of Genes and Genomes) database platform [1-3]. This database was chosen because it contains data on only full sequenced genomes, which is a requirement to search for the presence or absence of genes in specific species. In KEGG database, organisms are classified according to categories, without naming phyla, however for the clarity of the text we will use the term phyla when referring to any KEGG category inferior to domain. The information used in our study was that available by December 2015. We selected the first strain of each species and we always used the same strain for the analyses of all proteins or protein complexes. We performed BLASTp (default parameters were used) analyses of the different respiratory proteins using as templates sequences from *E. coli* K-12 MG1655, when no biochemical data was available. In the case of biochemically characterized proteins we used the respective amino acid sequence as query (Table IV.1.1). All amino acid sequences with E-value < 0.01 were accepted for analysis. Some proteins (or protein complexes) demanded further specific analyses to define the respective protein family:

a) In some Eukaryotic species we also performed protein analyses using BLAST (default parameters were used) at NCBI database [4] or analysis of other subunits included in the same protein complex. We noticed, in some cases, mitochondrial or chloroplastidial DNA was missing from KEGG database. These were the cases of Complex I, ubiquinol:cytochrome c oxidoreductase (bc_1 or complex III), plastoquinol:plastocyanin oxidoreductase (b_6f) and Heme-copper Oxygen reductases (HCO).

b) The amino acid sequence length was used as an additional criterion for selection in two cases. For type-1 copper proteins and cytochrome *c* only amino acid sequence between 67-300 and 51-202 amino acid residues, respectively, were accepted as positive hits.

c) The presence of the respective coding gene in a gene cluster and sometimes the gene cluster organization, were used as criteria for selection. Ferredoxin:methanophenazine oxidoreductase or ferredoxin:NAD^+ oxidoreductase (Rnf) family was identified by the presence of six genes (*rnfA*, *B*, *C*, *D*, *E* and *G*) in the gene cluster. Na^+ translocating NADH:quinone oxidoreductase (NQR) family was identified by the presence of six genes (*nqrA*, *B*, *C*, *D*, *E* and *F*) in the gene cluster. Both Rnf and NQR families' genes data sets were manually refined and compared since four subunits share high similarity.

Members of the Complex iron-sulfur molybdoenzyme (CISM) family were identified by the respective gene clusters. Gene clusters ordered by catalytic subunit, four clusters protein (FCP) and membrane attachment protein (MAP) subunit were used to identify

dimethyl sulfoxide (DMSO) reductase (DmsABC), Formate dehydrogenase-N (FdnGHI), Polysulfide reductase (PsrABC), Tetrathionate reductase (TtrABC), Thiosulfate reductase (PhsABS) and Sulfur reductase (SreABC).

Table IV.1.1: Respiratory enzymes, whose sequences were used as queries for pBLAST analyses.

Protein	Subunits	Species	KEGG ID	UniProt ID	PDB
Complex I family - Complex I	NuoA	*H. sapiens*	hsa:4537	P03897	-
Complex I family - Complex I	NuoA	*E. coli*	eco:b2288	P0AFC3	3RKO
Complex I family - Complex I	NuoA	*M. mazei*	mma:MM_2491	Q8PU58	-
Complex I family - Complex I	NuoH	*H. sapiens*	hsa:4535	P03886	-
Complex I family - Complex I	NuoH	*E. coli*	eco:b2282	P0AFD4	-
Complex I family - Complex I	NuoH	*M. mazei*	mma:MM_2487	Q8PU59	-
SDH	SdhA	*A. thaliana*	ath:AT5G66760	O82663	-
SDH	SdhB	*A. thaliana*	ath:AT5G40650	Q8LB02	-
SDH	SdhC	*E.coli*	eco:b0721	P69054	2WDQ, *et al*
SDH	SdhC	*B.subtilis*	bsu:BSU28450	P08064	-
SDH	SdhC	*G.sulfurreducens*	gsu:GSU1176	Q74DY9	-
SDH	SdhC	*R. marinus*	rmr:Rmar_0208	D0MD04	-
SDH	SdhC	*A. thaliana*	ath:AT5G09600	A8MSF5	-
SDH	SdhC	*H. pylori*	hpy:HP0193	O06912	-
SDH	SdhC	*O. japonica*	osa:4328125	Q6ZH92	-
SDH	SdhC	*F. Albicollis*	fab:101810785	UPI00035A00BD*	-
SDH	SdhC	*Syn. sp.*	syw:SYNW0592	Q7U8M4	-
ETF-QO	-	*E. coli*	eco:b0043	P68644	-
NDH-2	-	*E. coli*	eco:b1109	P00393	-
SQR	-	*A. aeolicus*	aae:aq_2186	O67931	3HYW, *et al*
DHOdh	-	*H. sapiens*	hsa:1723	Q02127	4IGH,*et al*
DHOdh	-	*E. coli*	eco:b0945	P0A7E1	1F76
L-Proline dehydrogenase	-	*E. anophelis*	eao:BD94_2425	A0A077EL88	3E1Q
DAAdh	-	*E. coli*	eco:b1189	P0A6J5	-
GPDH	GlpB	*E. coli*	eco:b2242	P13033	-
GPDH	GlpD	*E. coli*	eco:b3426	P13035	2QCU, *et al*
MQO	-	*E. coli*	eco:b2210	P33940	-
PQO	poxB	*E. coli*	eco:b0871	P07003	3EYA, 3EY9
Lqo	-	*E. coli*	eco:b2133	P06149	1F0X
mGDH	-	*G. oxydans*	gox:GOX0265	P27175	-
GLDH	SldB	*G. oxydans*	gox:GOX0855	Q70JP0	-
ADH	cytochrome *c*	*G.diazotrophicus*	gdi:GDI2041	A9HK15	-
ALDH	cytochrome *c*	*G. oxydans*	gox:GOX0585	Q5FTD3	-
SLDH	-	*G. oxydans*	gox:GOX2097	Q5FP63	-
GADH	γ	*P. aeruginosa*	pae:PA2264	Q9I1K9	-
PPOR	-	*E. coli*	eco:b3850	P0ACB4	-
Group-1 [NiFe] hydrogenases	HyaA	*E. coli*	eco:b0972	P69739	5ADU, *et al*
Group-1 [NiFe] hydrogenases	HyaB	*E. coli*	eco:b0973	P0ACD8	5ADU, *et al*
Group-1 [NiFe] hydrogenases	HyaC	*E. coli*	eco:b0974	P0AAM1	4GD3
Group-1 [NiFe] hydrogenases	HybA	*E. coli*	eco:b2996	P0AAJ8	-
Group-1 [NiFe] hydrogenases	HybB	*E. coli*	eco:b2995	P37180	-
Group-1 [NiFe] hydrogenases	HybC	*E. coli*	eco:b2994	P0ACE0	-
Group-1 [NiFe] hydrogenases	HybO	*E. coli*	eco:b2997	P69741	-
CISM family - Fdn-N	FdnG	*E. coli*	eco:b1474	P24183	1KQF, 1KQG
CISM family - Fdn-N	FdnH	*E. coli*	eco:b1475	P0AAJ3	1KQF, 1KQG
CISM family - Fdn-N	FdnI	*E. coli*	eco:b1476	P0AEK7	1KQF, 1KQG
QRC	QrcA	*D. vulgaris*	dvu:DVU4027	UPI0000E1DCD4*	-
QRC	QrcB	*D. vulgaris*	dvu:DVU0694	Q72E84	-
QRC	QrcC	*D. vulgaris*	dvu:DVU0693	Q72E85	-
QRC	QrcD	*D. vulgaris*	dvu:DVU0692	Q72E86	-
Nhc	nhcA	*D. desulfuricans*	dds:Ddes_2038	Q9RN68	1OFW, 1OFY
Nhc	nhcB	*D. desulfuricans*	dds:Ddes_2039	B8J3C5	-
Nhc	nhcC	*D. desulfuricans*	dds:Ddes_2040	B8J3C6	-
NQR	NqrA	*V. cholerae*	vch:VC2295	Q9KPS1	4U9O, 4U9Q
Menaquinone - MenB	-	*E. coli*	eco:b2262	P0ABU0	4I42, *et al*
Menaquinone - MenC	-	*E. coli*	eco:b2261	P29208	1R6W, *et al*
Menaquinone - MenD	-	*E. coli*	eco:b2264	P17109	3HWW, *et al*
Menaquinone - MenE	-	*E. coli*	eco:b2260	P37353	-
Menaquinone - MenF	-	*E. coli*	eco:b0593	P0AEJ2	3HWO
Menaquinone - MenH	-	*E. coli*	eco:b2263	P37355	4MYD, *et al*
Menaquinone - MenI	-	*E. coli*	eco:b1686	P77781	1VH5, *et al*
Menaquinone - MqnA	-	*C. jejuni*	cje:Cj1285c	Q0P8X2	-
Menaquinone - MqnC	-	*C. jejuni*	cje:Cj0462	Q0PB51	-
Menaquinone - MqnD	-	*C. jejuni*	cje:Cj1674	Q0P7V5	-
Menaquinone - MqnE	-	*C. jejuni*	cje:Cj1368	Q0P8P1	-

Protein	Subunits	Species	KEGG ID	UniProt ID	PDB
Ubiquinone - UbiA	-	*E. coli*	eco:b4040	P0AGK1	-
Ubiquinone - UbiD/X	-	*E. coli*	eco:b2311	P0AG03	-
Ubiquinone - UbiI	-	*E. coli*	eco:b2906	P25535	4K22
Ubiquinone - UbiH	-	*E. coli*	eco:b2907	P25534	-
Ubiquinone - UbiE	-	*E. coli*	eco:b3833	P0A887	-
Ubiquinone - UbiF	-	*E. coli*	eco:b0662	P75728	-
Ubiquinone - UbiG	-	*E. coli*	eco:b2232	P17993	4KDR, 4KDC
Ubiquinone - Coq2	-	*S. cerevisiae*	sce:YNR041C	P32378	-
Ubiquinone - Coq3	-	*S. cerevisiae*	sce:YOL096C	P27680	-
Ubiquinone - Coq5	-	*S. cerevisiae*	sce:YML110C	P49017	4OBX, 4OBW
Ubiquinone - Coq6	-	*S. cerevisiae*	sce:YGR255C	P53318	-
Ubiquinone - Coq7	-	*S. cerevisiae*	sce:YOR125C	P41735	-
Plastoquinone/Tocopherol - PDS1	-	*A.thaliana*	ath:AT1G06570	F4IDP2	1TFZ, *et al*
tails - Phytyl-PP	-	*A. thaliana*	ath:AT1G74470	Q9CA67	-
tails - hexaprenyl-PP	-	*S. cerevisiae*	sce:YBR003W	P18900	-
tails - Octaprenyl-PP	-	*E. coli*	eco:b3187	P0AD57	3WJK, *et al*
tails - Nonaprenyl-PP	-	*A.thaliana*	ath:AT1G17050	Q76FS5	-
tails - Decaprenyl-PP	-	*H. sapiens*	hsa:23590	Q5T2R2	-
bc_1	cytochrome *b*	*R. sphaeroides*	rsp:RSP_1395	Q3IY10	2QJY, *et al*
bc_1	cytochrome *b*	*S. acidocaldariu*	sai:Saci_1861	Q4J7R5	-
b_6f	cytochrome *f*	*Cya. sp.*	cyt:cce_2959	B1WVX4	-
ACIII	ActA	*R. marinus*	rmr:Rmar_0221	D0MDD4	-
ACIII	ActB	*R. marinus*	rmr:Rmar_0222	D0MDD5	-
ACIII	ActC	*R. marinus*	rmr:Rmar_0223	D0MDD6	-
ACIII	ActD	*R. marinus*	rmr:Rmar_0224	D0MDD7	-
ACIII	ActE	*R. marinus*	rmr:Rmar_0225	D0MDD8	-
ACIII	ActF	*R. marinus*	rmr:Rmar_0226	D0MDD9	-
HDR	hdrE	*M. mazei*	mma:MM_1843	Q8PVW4	-
Qmo	qmoA	*D. desulfuricans*	dds:Ddes_2127	B8J3X8	-
Qmo	qmoB	*D. desulfuricans*	dds:Ddes_2126	B8J3X7	-
Qmo	qmoC	*D. desulfuricans*	dds:Ddes_2125	B8J3X6	-
Cytochrome *c*	-	*R. sphaeroides*	rsp:RSP_0296	Q3J164	1CXC, *et al*
Cytochrome *c*	-	*G. sulfurreducens*	gsu:GSU0068	Q74H25	-
Cytochrome *c*	-	*S. elongatus*	syc:syc0089_d	P07497	-
Cytochrome *c*	-	*Cya sp.*	cyt:cce_0589	A1KYG5	-
Cytochrome *c*	-	*R. marinus*	rmr:Rmar_1369	D0MIF0	-
Cytochrome *c*	-	*C. tepidum*	cte:CT0188	Q8KFY2	-
Cytochrome *c*	-	*T. thermophilus*	tth:TTC0872	Q72JA7	-
Cytochrome *c*	-	*A. pernix*	ape:APE_1719.1	Q9YB78	-
HiPIP	-	*R. marinus*	rmr:Rmar_1811	D0MJN7	3H31
Type I Cooper proteins -Amicyanin	-	*P. denitrificans*	pde:Pden_4734	A1BBA1	2OV0, *et al*
Type I Cooper proteins -Auracyanin	-	*C. aurantiacus*	cau:Caur_1950	P27197	1QHQ, 1OV8
Type I Cooper proteins -Azurin	-	*A. xylosoxidans*	axy:AXYL_04546	E3HH47	1AZC, *et al*
Type I Cooper proteins -Halocyanin	-	*N. pharaonis*	-	P39442	1N3D (model)
Type I Cooper proteins -Nitrosocyanin	-	*N. europaea*	neu:NE0143	Q820S6	1IBY, *et al*
Type I Cooper proteins -Plastocyanin	-	*S. elongatus*	syc:syc0461_c	Q5N4W8	1BXV, 1BXU
Type I Cooper proteins -Rusticyanin	-	*A. ferrooxidans*	afr:AFE_3146	B7JAQ0	2CAL, *et al*
Type I Cooper proteins -Stellacyanin	-	*C.sativus*	csv:101209347	Q96403	1JER
Type I Cooper proteins -Sulfocyanin	-	*S. acidocaldarius*	sai:Saci_2262	Q53765	-
Type I Cooper proteins -Uclacyanin	-	*A.thaliana*	ath:AT2G32300	O82081	-
HCO - A1, aa_3 cytochrome c oxidase	I	*P. denitrificans*	pde:Pden_1938	A1B3D7	3HB3, *et al*
HCO - A2, caa_3 cytochrome c oxidase	I	*T. thermophilus*	tth:TTC1671	Q72H24	2YEV
HCO - B, ba_3 cytochrome c oxidase	I	*T. thermophilus*	tth:TTC0770	Q72JK2	3S8F, *et al*
HCO - C, cbb_3 cytochrome c oxidase	I	*P. stutzeri*	psa:PST_1843	A4VKL6	5DJQ, 3MK7
HCO - A1, bo_3 quinol oxidase	I	*E. coli*	eco:b0431	P0ABI8	1FFT
HCO - B, *aa* 3 quinol oxidase	soxB	*S. solfataricus*	sso:SSO2657	Q97VG9	-
HCO - A1, *aa* 3 quinol oxidase	soxM	*S. solfataricus*	sso:SSO2973	Q97UN0	-
bd quinol oxidase	I	*E. coli*	eco:b0978	P26459	-
AOX	-	*T.brucei*	tbr:Tb10.6k15.3640	Q38AQ4	3W54, *et al*
qNOR	NorB	*G. stearothermophilus*	gse:GT50_13445	A0A0K2HAS5	3AYF
cNOR	NorB	*P. aeruginosa*	pae:PA0524	Q59647	3O0R, *et al*
Dsr	DsrM	*D. vulgaris*	dvu:DVU1290	Q72CJ3	-
Dsr	DsrK	*D. vulgaris*	dvu:DVU1289	Q72CJ4	-
Dsr	DsrJ	*D. vulgaris*	dvu:DVU1288	Q72CJ5	-
Dsr	DsrO	*D. vulgaris*	dvu:DVU1287	Q72CJ6	-
Dsr	DsrP	*D. vulgaris*	dvu:DVU1286	Q72CJ7	-
CISM family - DMSO reductase	DmsA	*E. coli*	eco:b0894	P18775	-
CISM family - DMSO reductase	DmsB	*E. coli*	eco:b0895	P18776	-
CISM family - DMSO reductase	DmsC	*E. coli*	eco:b0896	P18777	-

Protein	Subunits	Species	KEGG ID	UniProt ID	PDB
CISM family - DMSO reductase	DmsA	*E. coli*	eco:b0894	P18775	-
CISM family - DMSO reductase	DmsB	*E. coli*	eco:b0895	P18776	-
CISM family - DMSO reductase	DmsC	*E. coli*	eco:b0896	P18777	-
CISM family - Nitrate reductase	NarG	*E. coli*	eco:b1224	P09152	3EGW, *et al*
CISM family - Nitrate reductase	NarH	*E. coli*	eco:b1225	P11349	3EGW, *et al*
CISM family - Nitrate reductase	NarJ	*E. coli*	eco:b1226	P0AF26	-
CISM family - Nitrate reductase	NarI	*E. coli*	eco:b1227	P11350	3EGW, *et al*
Periplasmic Nitrate reductase	NapA	*E. coli*	eco:b2206	P33937	2NYA, 2PQ4
Periplasmic Nitrate reductase	NapB	*E. coli*	eco:b2203	P0ABL3	-
Periplasmic Nitrate reductase	NapC	*E. coli*	eco:b2202	P0ABL5	-
Periplasmic Nitrate reductase	NapG	*E. coli*	eco:b2205	P0AAL3	-
Periplasmic Nitrate reductase	NapH	*E. coli*	eco:b2204	P33934	-
Cytocrome *c* nitrite reductase	NrfA	*E. coli*	eco:b4070	P0ABK9	2RDZ, *et al*
Cytocrome *c* nitrite reductase	NrfB	*E. coli*	eco:b4071	P0ABL1	2OZY, 2P0B
Cytocrome *c* nitrite reductase	NrfC	*E. coli*	eco:b4072	P0AAK7	-
Cytocrome *c* nitrite reductase	NrfD	*E. coli*	eco:b4073	P32709	-
Cytocrome *c* nitrite reductase	NrfA	*W. succinogenes*	wsu:WS0969	Q9S1E5	3BNJ, *et al*
Cytocrome *c* nitrite reductase	NrfH	*W. succinogenes*	wsu:WS0970	Q9S1E6	-
AMO/pMMO monooxygenase	A	*N. europaea*	nmr:Nmar_1500	A9A4U2	-
AMO/pMMO monooxygenase	A	*N. europaea*	pde:Pden_0324	A1AYU4	-
AMO/pMMO monooxygenase	A	*N. europaea*	neu:NE2063	Q04507	-
CISM family - Polysulfide reductase	PsrA	*W. succinogenes*	wsu:ws0116	P31075	-
CISM family - Polysulfide reductase	PsrB	*W. succinogenes*	wsu:ws0117	P31076	-
CISM family - Polysulfide reductase	PsrC	*W. succinogenes*	wsu:ws0118	P31077	-
CISM family - Sulfur reducing complex	SreA	*A. aeolicus*	aae:aq_1234	O67280	-
CISM family - Sulfur reducing complex	SreB	*A. aeolicus*	aae:aq_1232	O67279	-
CISM family - Sulfur reducing complex	SreC	*A. aeolicus*	aae:aq_1231	O67278	-
CISM family - Tetrathionate reductase	TtraA	*B. parapertussi*	bpa:BPP1028	Q7WBH5	-
CISM family - Tetrathionate reductase	TtraB	*B. parapertussi*	bpa:BPP1030	Q7WBH3	-
CISM family - Tetrathionate reductase	TtraC	*B. parapertussi*	bpa:BPP1029	Q7WBH4	-
CISM family - Thiosulfate reductase	PhsA	*S. enterica*	stm:STM2065	P37600	-
CISM family - Thiosulfate reductase	PhsB	*S. enterica*	stm:STM2064	P0A1I1	-
CISM family - Thiosulfate reductase	PhsC	*S. enterica*	stm:STM2063	P37602	-
TMAO reductase	TorA	*E. coli*	eco: b0997	P33225	-
TMAO reductase	TorC	*E. coli*	eco:b0996	P33226	-
TMAO reductase	TorD	*R. sphaeroides*	rsp:RSP_3048	Q3IXS0	1EU1
Complex I family - Ech	EchB	*M.mazei*	mma:MM_2321	Q8PUL3	-
Complex I family - Ehr	EhrB	*M.mazei*	mma:MM_1060	Q8PY06	-
Complex I family - Hyc	HycC	*E. coli*	eco:b2722	P16430	-
Complex I family - Hyf	HyfC	*E. coli*	eco:b2483	P77858	-
Complex I family - Mbh	MbhM	*P. furiosus*	pfu:PF1435	Q8U0Z5	-
Complex I family - Mbx	MbxM	*P. furiosus*	pfu:PF1445	Q8U0Y6	-
Sodium ion-translocating decarboxylases	β	*K.pneumoniae*	kpm:KPHS_07400	A0A0H3GJ14	-
Sodium ion-translocating decarboxylases	β	*A. fermentans*	afn:Acfer_1834	Q9ZAA6	-
Sodium ion-translocating decarboxylases	β	*Des. sp.*	des:DSOUD_0389	A0A0M4CZ24	-
Sodium ion-translocating decarboxylases	β	*R. palustris*	rpb:RPB_3300	Q2IUW4	-
Mtr	MtrC	*M. mazei*	mma:MM_1545	O59638	-
mPPase	-	*T. maritima*	tma:TM0174	Q9S5X0	4AV3, 4AV6
Rnf	RnfB	*V. cholerae*	vch:VC1016	Q9KT87	-
ATP synthase - F-type	γ	*S. cerevisiae*	sce:Q0085	P00854	-
ATP synthase - F-type	γ	*Es. coli*	eco:b3738	P0AB98	1C17
ATP synthase - A-type	D	*S. cerevisiae*	sce:YEL051W	P32610	4RND, *et al*
ATP synthase - V-type	E	*M. mazei*	mma:MM_0778	Q60188	-
Pnt	β	*T. thermophilus*	tth:TTC1778	Q72GS0	4J1T, *et al*

* Protein with only UniParc ID available.

For the respiratory nitrate reductase (NarGHI) we also included *narJ* (gene coding for a specific chaperone) in our analysis. For quinol:trimethylamine N-oxide oxidoreductase (TMAO) reductase and in order to distinguish it from DorCA, a gene coding for a chaperone was also included in the analysis. *torA* and *torC* are consecutive while *dorC* and *dorA* have the referred gene between them. Due to the lack of homology between the membrane subunits of [NiFe] hydrogenases,

searches for gene clusters coding for this enzyme were performed using two different templates: *E. coli* Hydrogenase-1 (Hya) and Hydrogenase-2 (Hyd). Several gene clusters have been described to code for Periplasmic Nitrate reductase (NAP) and therefore we searched for all the gene clusters *napABC*, *napABHG* and *napABCGH*. The order of the genes in the cluster was irrelevant. For cytochrome *c* nitrite reductase, NrfA, we investigated the presence of *nrfAH* and *nrfABCD* gene clusters.

Alternative complex III (ACIII) was identified by the presence of the gene cluster formed by *actABCDEF* or *actAB1B2CDEF*. Quinone reductase complex (QRC) was identified by the genes *qrcABCD*. The search of Qmo complexes (Quinone-interacting membrane-bound oxidoreductase) was achieved by gene cluster identification. Since this protein complex was first characterized in *Desulfovibrio desulfuricans* ATCC 27774 [5], we used the protein from this organism as the initial template. As all the three subunits have homology with proteins from other complexes none of them could be used as a defining template to retrieve the remaining Qmo family. To do so we blasted QmoA, QmoB and QmoC and we only considered as positive results gene clusters containing these three subunits.

A similar approach was performed for Nhc complex (nine-heme cytochrome). To search for genes coding for this protein family in the KEGG database, we considered as positive results, the gene clusters that contained three genes, homologous to *NhcA*, *NhcB* and *NhcC* from *D. desulfuricans* [6]. *NhcD* was not considered since it is too small to generate reliable results. Dissimilatory sulfite-reductase (Dsr)

complex was identified by the presence of gene clusters encoding subunits *dsrM* and *dsrK*. The gene clusters that encode the five subunits (*dsrM*, *dsrK*, *dsrJ*, *dsrO* and *dsrP*) were also identified. This consideration was based on the described *D. desulfuricans* gene cluster [7].

For Complex I and related proteins, we performed before their respective taxonomic analysis using KEGG´s information available by April 2012 (section II.1); we now identified the new species deposited at KEGG database and updated the previous taxonomic profile [8]. In this case, we performed BLASTp using the subunit NuoH from Complex I and homologous subunits from the other complexes belonging to the family. In order to identify complex I among Complex I family, we performed BLASTp using NuoA from Complex I, a specific subunit for this enzyme (Table IV.1.1).

d) In the case of protoporphyrinogen IX:quinone oxidoreductase (PPOR), D-amino acid:quinone oxidoreductase (DAADH), dihidroorotate:quinone oxidoreductase (DHODH) and malate:quinone oxidoreductase (MQO) we performed amino acid sequence alignments using PROMALS3D [9] and respective Neighbor-joining (NJ) [10] trees with 10 replicates of bootstrap and 3 generated seeds.

We previously performed a taxonomic analysis for NDH-2 and sulfide:quinone oxidoreductase (SQR) families with KEGG´s information available by September 2014 (section III.1) and now we identified the new species deposited in KEGG database to update the previous taxonomic profile. Thus, we aligned the new amino acid

sequences with previously identified amino acid sequences of NDH-2 and SQR using ClustalX 2.1 [10]. We defined as SQR the proteins that present the typical cysteine or methionine residues that are part of the catalytic center (approximately 10 % of the sequences present a methionine in the place of the second cysteine in the redox active pair).

The taxonomic profiling of the species with the protein family of interest was performed using a homemade script on Microsoft Office Excel using Visual Basic for Applications (VBA) programming language.

IV.1.2.2 Quinone profiling

We considered that a species is able to synthesize a certain type of quinone if it contains the genes coding for enzymes involved in the synthesis of the core aromatic ring of the respective quinone (Table IV.1.1) according to the different biosynthetic pathways indicated at KEGG database (ubiquinone and other terpenoid-quinone biosynthesis, map00130).

We considered the presence of ubiquinone if simultaneous presence of the seven genes *ubiA, ubiD/X, ubiE, ubiF, ubiG, ubiI, ubiH* or of the five genes *coq2*, *coq3*, *coq5*, *coq6*, *coq7* is observed. The presence of menaquinone is attested by the observation of the simultaneous presence of the four genes *MqnA, MqnC, MqnD* and *MqnE* (we did not consider *MqnB* because the function of the encoded protein may be performed by alternative enzymes), or of the seven genes *menB*, *menC*, *menD*, *menH*, *menI*,

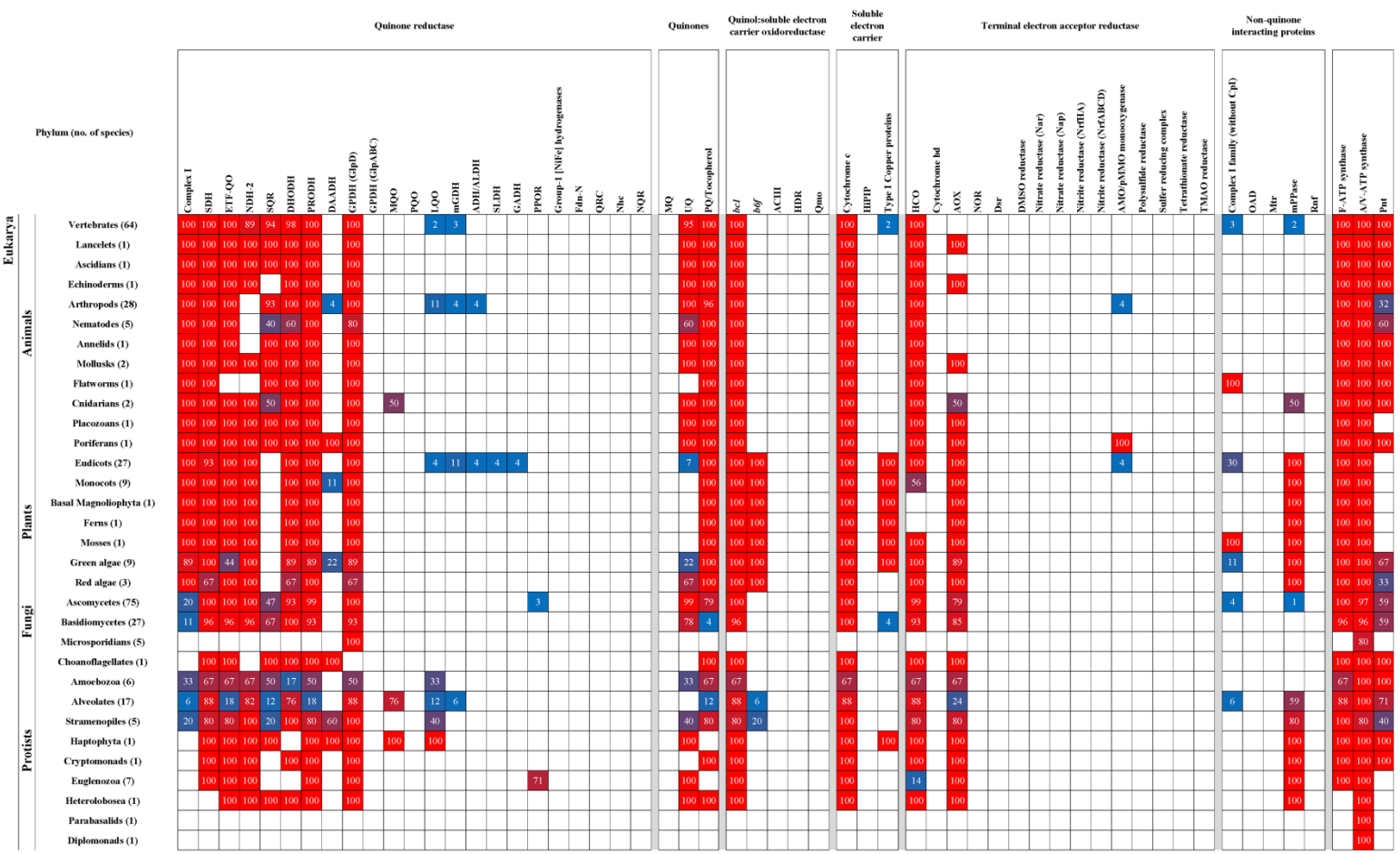
menE, *menF*. We concluded platoquinone/tocopherol is synthesized if the gene coding for 4-hydroxyphenylpyruvate dioxygenase (EC: 1.13.11.27) is present in the genome of the respective organisms.

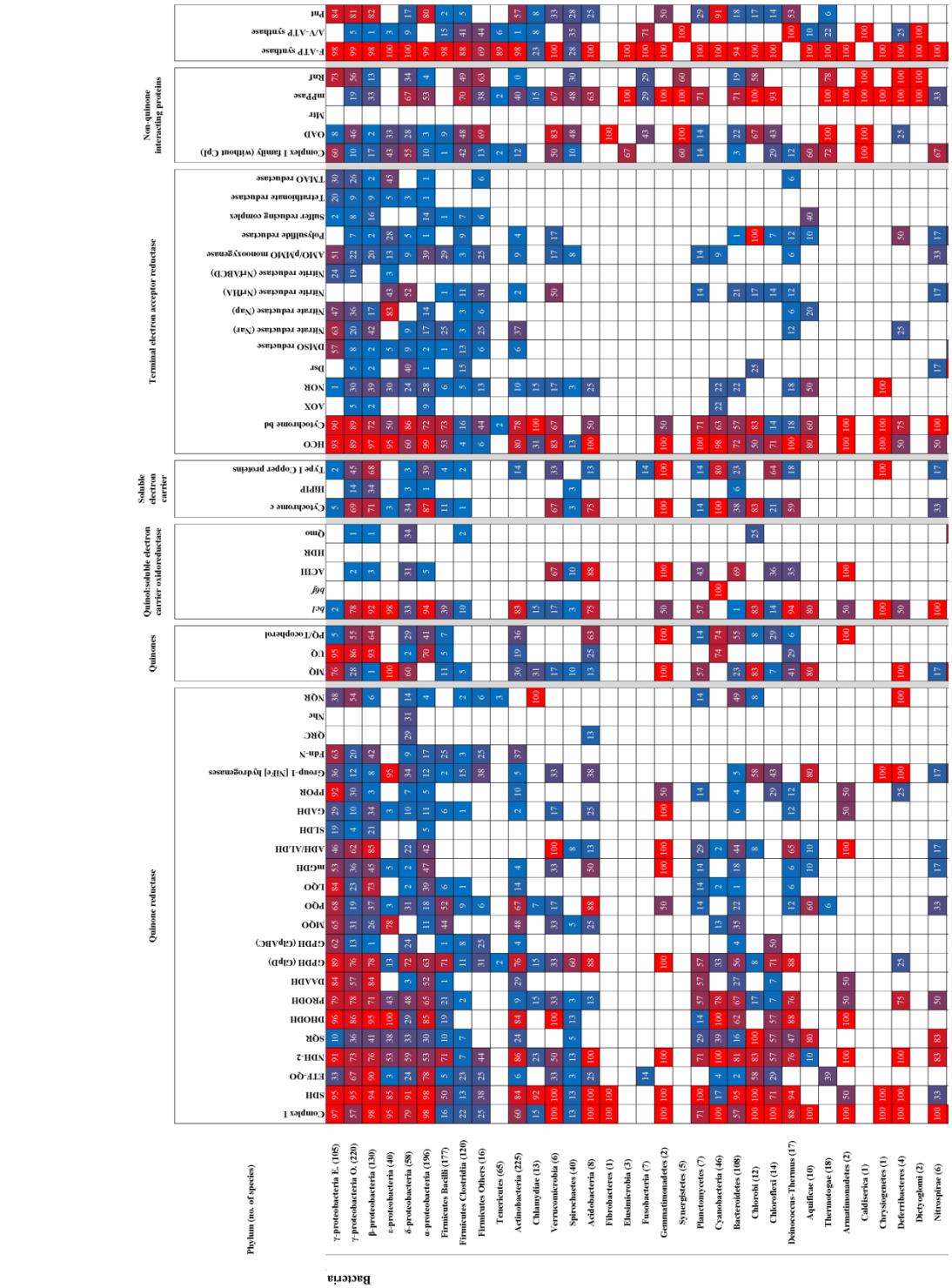
Moreover, we also identified the genes involved in the synthesis of isoprenoid tails (phytyl-PP, hexaprenyl-PP, octaprenyl-PP, nonaprenyl-PP and decaprenyl-PP) (Supplementary Figure IV.1.1).

The absence of a positive hit for quinone synthesis means the respective organism does not produce quinones by the analyzed pathways, but does not exclude the possibility that organisms may synthetize an alternative quinone system or the discussed quinones by a different pathway, such as the case of the methanophenazine biosynthetic pathway which is still not totally known. We further investigated the presence of quinones using the reports in the literature for the respective biochemical characterization (Figure IV.1.1).

IV.1.3 Results and Discussion

In this work we explored the diversity of membrane respiratory chains and the presence of the different enzyme complexes in the several phyla of life. We performed taxonomic profiles of the different respiratory proteins and protein complexes evaluating the presence of their respective coding genes in all species deposited in KEGG database (Figure IV.1.1). We did not aim to present an exhaustive description of proteins involved in membrane-bound respiratory chains, but to bring focus to less discussed enzymes, which in many


Phylum (no. of species)
Quinone reductase
Complex I
SDH
ETF-QO
NDH-2
SQR
DHODH
PRODH
DAADH
GPDH (GlpD)
GPDH (GlpABC)
MQO
PQO
LQO
mGDH
ADH/ALDH
SLDH
GADH
PPOR
Group-1 [NiFe] hydrogenases
Fdn-N
QRC
Nhc
NQR
Quinones
MQ
UQ
PQ/Tocopherol
Quinol:soluble electron carrier oxidoreductase
bc1
bd/
ACIII
HDR
Quo
Soluble electron carrier
Cytochrome c
HiPIP
Type I Copper proteins
Terminal electron acceptor reductase
HCO
Cytochrome bd
AOX
NOR
Dsr
DMSO reductase
Nitrate reductase (Nar)
Nitrate reductase (Nap)
Nitrite reductase (NrfHA)
Nitrite reductase (NrfABCD)
AMO/pMMO monooxygenase
Polysulfide reductase
Sulfer reducing complex
Tetrathionate reductase
TMAO reductase
Non-quinone interacting proteins
Complex I family (without CpI)
OAD
Mtr
mPPase
Rnf
F-ATP synthase
AAV-ATP synthase
Pnt
Eukarya
Animals
Plants
Fungi
Protists
Vertebrates (64)
Lancelets (1)
Ascidians (1)
Echinoderms (1)
Arthropods (28)
Nematodes (5)
Annelids (1)
Mollusks (2)
Flatworms (1)
Cnidarians (2)
Placozoans (1)
Poriferans (1)
Eudicots (27)
Monocots (9)
Basal Magnoliophyta (1)
Ferns (1)
Mosses (1)
Green algae (9)
Red algae (3)
Ascomycetes (75)
Basidiomycetes (27)
Microsporidians (5)
Choanoflagellates (1)
Amoebozoa (6)
Alveolates (17)
Stramenopiles (5)
Haptophyta (1)
Cryptomonads (1)
Euglenozoa (7)
Heterolobosea (1)
Parabasalids (1)
Diplomonads (1)


Bacteria
Phylum (no. of species)
Quinone reductase
Complex I
SDH
ETF-QO
ZDH-2
SQR
DHODH
PRODH
DAADH
GlpDH (GlpD)
GlpDH (GlpABC)
NQO
PQO
LQO
mGDH
ADH/ALDH
SLDH
GADH
PFOR
Group-1 [NiFe] hydrogenases
Fdh-N
QRC
Nhe
NQR
Quinones
MQ
UQ
PQ/Tocopherol
Quinol:soluble electron carrier oxidoreductase
bc1
bf
ACIII
HDR
Qmo
Soluble electron carrier
Cytochrome c
HiPIP
Type I Copper proteins
Terminal electron acceptor reductase
HCO
Cytochrome bd
AOX
NOR
Dsr
DMSO reductase
Nitrate reductase (Nar)
Nitrate reductase (Nap)
Nitrite reductase (NrfHA)
Nitrite reductase (NrfABCD)
AMO/pMMO monooxygenase
Polysulfide reductase
Sulfur reducing complex
Tetrathionate reductase
TMAO reductase
Non-quinone interacting proteins
Complex I family (without CpI)
OAD
Mbr
mnPPase
Rnf
F-ATP synthase
A/V-ATP synthase
Pnt

Archaea
Phylum (no. of species)
Thermodesulfobacteria (3)
Unclassified (9)
Euryarchaeota (118)
Crenarchaeota (40)
Thaumarchaeota (9)
Nanoarchaeota (1)
Korarchaeota (1)
Lokiarchaeota (1)
Unclassified (3)
Quinone reductase
Complex I
SDH
ETF-QO
NDH-2
SQR
DHODH
PRODH
DAADH
GPDH (GlpD)
G3PDH (GlpABC)
MQO
PQO
LQO
mGDH
ADH/ALDH
SLDH
GADH
PFOR
Group-1 [NiFe] hydrogenases
Fdh-N
QRC
Nhc
NQR
Quinones
MQ
UQ
PQ/Tocopherol
Quinol-soluble electron carrier oxidoreductase
bc-f
bbj
ACIII
HDR
Qmo
Soluble electron carrier
Cytochrome c
HiPIP
Type I Copper proteins
Terminal electron acceptor reductase
HCO
Cytochrome bd
AOX
NOR
Dsr
DMSO reductase
Nitrate reductase (Nar)
Nitrate reductase (Nap)
Nitrite reductase (NrfHA)
Nitrite reductase (NrfABCD)
AMO/pMMO monooxygenase
Polysulfide reductase
Sulfer reducing complex
Tetrathionate reductase
TMAO reductase
Non-quinone interacting proteins
Complex I family (without CpI)
OAD
Mfr
mrPPase
Rnf
F-ATP synthase
A/V-ATP synthase
Pnt

Figure IV.1.1: Taxonomic distribution of membrane-bound **proteins and protein complexes involved in respiratory chains and in the biosynthesis pathways of different quinones**. The frequency of genes coding for each enzyme per phyla is represented by a color gradient from blue (1 %) to red (100 %). White color corresponds to the absence of the gene coding for the enzyme. Quinone reductases: Complex I (NADH:quinone oxidoreductase), including the Complex I-like enzymes (bacterial 11 subunit complexes or archaeal complexes); SDH (succinate:quinone oxidoreductase); ETF-QO (Electron transfer flavoprotein:quinone oxidoreductase); NDH-2 (Type-2 NADH:quinone oxidoreductase); SQR (sulfide:quinone oxidoreductase); DHODH (dihidroorotate:quinone oxidoreductase); PRODH (L-proline:quinone oxidoreductase); DAADH (D-amino acid:quinone oxidoreductase); G3PDH-GlpD (aerobic Glycerol-3-phosphate:quinone oxidoreductase); G3PDH-GlpABC (anaerobic Glycerol-3-phosphate:quinone oxidoreductase); MQO (Malate:quinone oxidoreductase); PQO (Pyruvate:quinone oxidoreductase); LQO (Lactate:quinone oxidoreductase); mGDH and GLDH (Glucose:quinone oxidoreductase and glycerol:quinone oxidoreductase); ADH/ALDH (Alcohol/Aldehyde:quinone oxidoreductases); SLDH (D-Sorbitol:quinone oxidoreductase); GADH (D-Gluconate:quinone oxidoreductase); PPOR (Protoporphyrinogen IX:quinone oxidoreductase); Group-1 [NiFe] Hydrogenase (hydrogen:quinone oxidoreductase); Fdn-N (formate:quinone oxidoreductase); QRC (cytochrome c_3:quinone oxidoreductase); Nhc (cytochrome:quinone oxidoreductase); NQR (Na^+ translocating NADH:quinone oxidoreductase); Quinones: MQ (Menaquinone); UQ (Ubiquinone); PQ (Plastoquinone); Quinol:soluble electron carrier oxidoreductase: bc_1 (cytochrome bc_1 complex); b_6f (cytochrome b_6f complex); ACIII (Alternative complex III); HDR (methanophenazine:heterodisulfide oxidoreductase); Qmo (quinol:electron acceptor oxidoreductase); Soluble electron carriers: Cytochromes *c*; HiPIP (high potential iron-sulfur protein); type 1 Copper Proteins (Azurin, Amicyanin, Plastocyanin, Pseudoazurin, Rusticyanin, Auracyanin, Stellacyanin, Plantacyanin, Uclacyanin, Halocyanin, Sulfocyanin and Nitrosocyanin); Terminal electron acceptor reductases: HCO (Heme-copper Oxygen reductases); cytochrome *bd* oxidase (Quinol:oxygen oxidoreductase); AOX (Alternative quinol:oxygen oxidoreductase); NOR (nitric oxide reductase); Dsr (Dissimilatory sulfite-reductase complex); DmsABC (DMSO reductase); Nar (Quinol:nitrate oxidoreductase); Nap (Periplasmic quinol:nitrate oxidoreductase); NrfHA (Quinol:nitrite oxidoreductase); NrfABCD (Quinol:nitrite oxidoreductase); AMO/pMMO (Ammonia monooxygenase/Particulate methane monooxygenase family); PsrABC (Polysulfide reductase); SreABC (Sulfur redutase complex); TtrABC (Tetrathionate reductase), TMAO reductase (Quinol:trimethylamine N-oxide oxidoreductases); Non quinone interacting ion translocating oxidoreductases: Complex I family (Hyc, Hyf, Mbh, Ech, Eha, Ehb, Mbx and Ehr); OAD (Na^+-translocating decarboxylases); Mtr (Methyltetrahydromethanopterin:coenzyme M methyltransferase); mPPase (Membrane-bound pyrophosphatase); Rnf (Ferredoxin:methanophenazine/NAD^+ oxidoreductase); Others: F-ATP synthase; A/V-ATP synthase; Pnt (membrane-bound transhydrogenase).

cases are not described as part of those chains.

We described in our study membrane-bound enzymes or enzyme complexes involved in any respiratory process. We considered as respiratory processes all those that contribute to the establishment of the transmembrane difference of the electrochemical potential and thus allow the synthesis of ATP by ATP synthase. This distinguishes from a fermentative process, which involves the production of ATP only at substrate level phosphorylation.

We analyzed all proteins that reduce membrane quinones and proteins that oxidize quinols, in addition to membrane complexes involved in the reaction with final electron acceptors and those that perform redox or decarboxylation driven ion translocation. For consistency, we also investigated the electron carriers that make functional connection between the membrane complexes, quinones or soluble proteins. As ATP synthase and transhydrogenase may also establish membrane potential, under certain conditions, they were also included in our study.

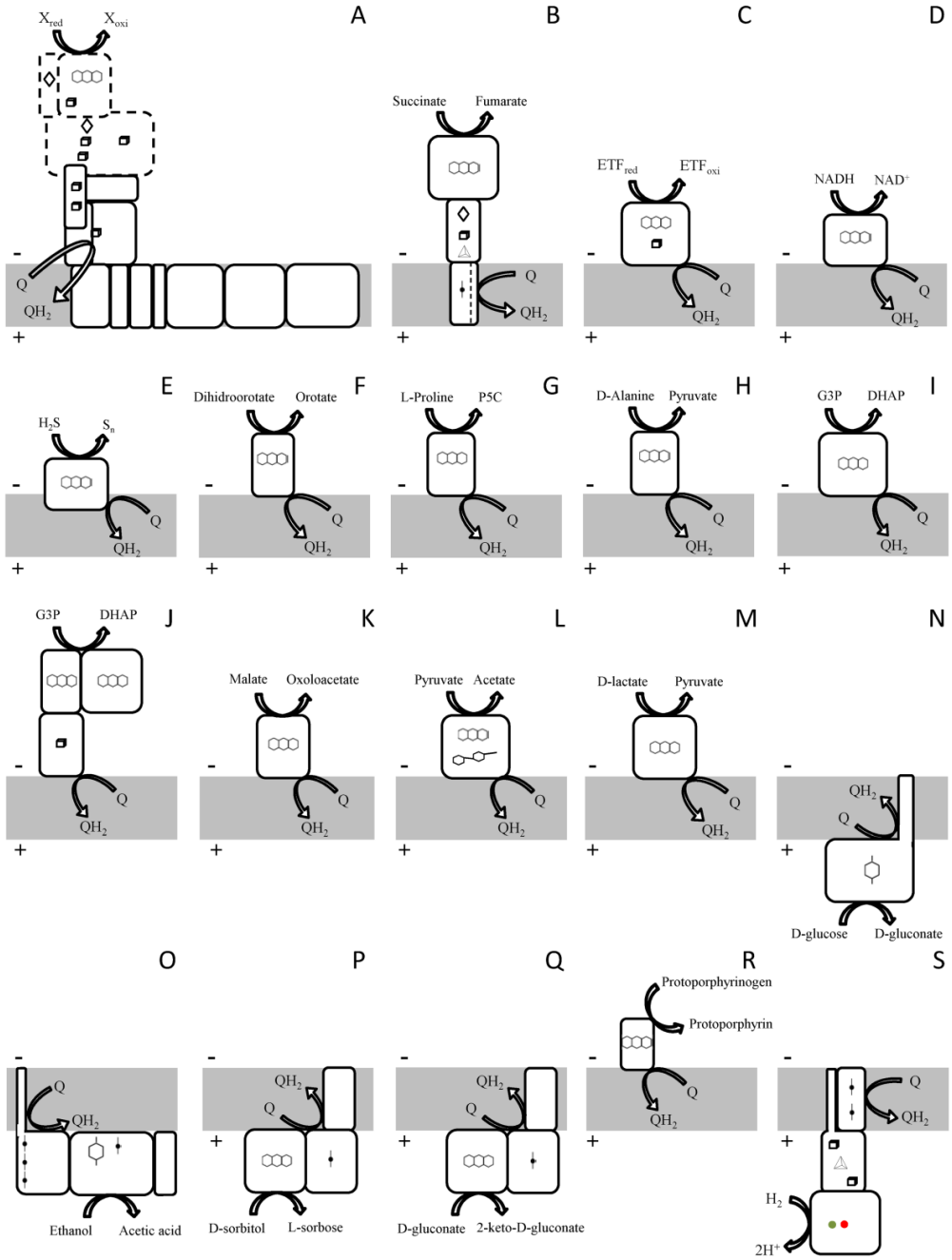
Quinone reductases are the first enzymes feeding electrons to membrane-bound respiratory chains. Classically only Complex I and succinate dehydrogenase (Complex II or SDH) are viewed as this type of enzymes; however, the plethora of enzymes able to oxidize several metabolites with quinone reduction is quite large: Electron transfer flavoprotein:quinone oxidoreductase (ETF-QO), NDH-2, SQR, DHODH, L-Proline:quinone oxidoreductase (PRODH), DAAD, Glycerol-3-phosphate:quinone oxidoreductase (G3PDH), MQO, Pyruvate:quinone oxidoreductase (PQO), Lactate:quinone oxidoreductase (LQO),

Glucose:quinone oxidoreductase (mGDH), Alcohol/Aldehyde:quinone oxidoreductases (ADH/ALDH), Sorbitol:quinone oxidoreductase (SLDH), Gluconate:quinone oxidoreductase (GADH), PPOR, Group-1 [NiFe] hydrogenases, Fdn-N, QRC, Nhc, Thiosulfite:quinone oxidoreductase (TQO) and NQR (Figure IV.1.1 and Figure IV.1.2).

As lipophilic electron carriers we analyzed the taxonomic distribution of ubiquinone, plastoquinone/tocopherol and menaquinone taking into account the presence of the enzymes involved in the respective biosynthesis pathways (Figure 1.1.3, Figure IV.1.1 and Supplementary Figure IV.1.1).

In the mitochondrial respiratory chain, quinol:cytochrome *c* oxidoreductase activity is performed by cytochrome bc_1 and in the chloroplast the same activity is carried out by cytochrome b_6f. In fact for long time that activity was considered to be exclusively done by the members of the cytochrome bc_1/b_6f complex family. This concept was changed upon identification and characterization of the so-called Alternative Complex III, particularly in organisms that lack bc_1/b_6f complexes, but which ought to have quinol:electron acceptor oxidoreductase activity due to the presence of soluble electron carrier:oxygen oxidoreductases. We also included in quinol:soluble electron carrier oxidoreductases the methanophenazine: heterodisulfide oxidoreductase (HDR) and Qmo complex (Figure IV.1.1 and Figure IV.1.3).

Genes coding for soluble electron carriers were also analyzed (cytochrome *c*, HiPIP and type 1 copper proteins). The taxonomic distribution for soluble electron carriers is presented in Figure IV.1.1.


A
X_{red}
X_{oxi}
Q
QH_2
B
Succinate
Fumarate
Q
QH_2
C
ETF_{red}
ETF_{oxi}
Q
QH_2
D
NADH
NAD^+
Q
QH_2
E
H_2S
S_n
Q
QH_2
F
Dihidroorotate
Orotate
Q
QH_2
G
L-Proline
P5C
Q
QH_2
H
D-Alanine
Pyruvate
Q
QH_2
I
G3P
DHAP
Q
QH_2
J
G3P
DHAP
Q
QH_2
K
Malate
Oxoloacetate
Q
QH_2
L
Pyruvate
Acetate
Q
QH_2
M
D-lactate
Pyruvate
Q
QH_2
N
QH_2
Q
D-glucose
D-gluconate
O
Q
QH_2
Ethanol
Acetic acid
P
QH_2
Q
D-sorbitol
L-sorbose
Q
QH_2
Q
QH_2
D-gluconate
2-keto-D-gluconate
R
Protoporphyrinogen
Protoporphyrin
Q
QH_2
S
Q
QH_2
H_2
$2H^+$

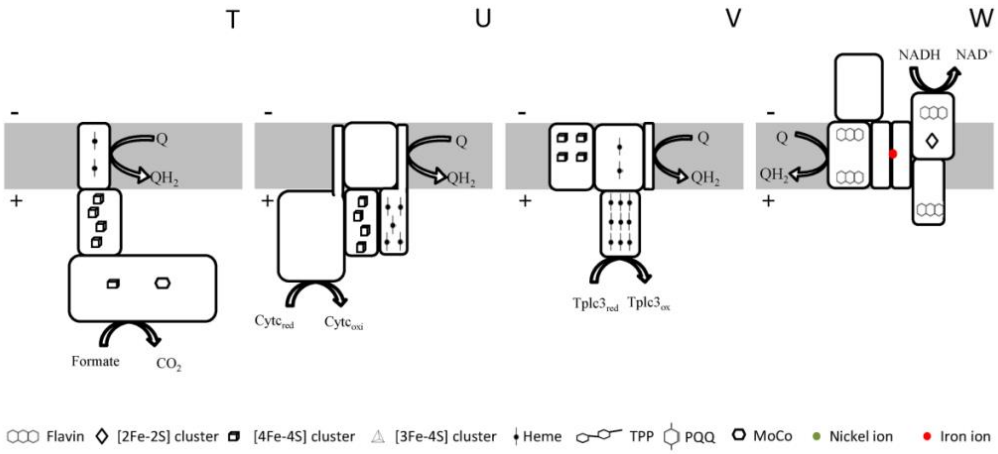


Figure IV.1.2: Quinone reductases present in respiratory chains. Schematic representation of proteins or protein complexes of respiratory chains involved in quinone reduction. In case of homo-oligomers, only monomers are represented. Protein representation is proportional to its MM. **A** – Complex I (nuoA to N) with different input models (dashed subunits are not present in all enzymes) allowing interaction with different substrates (e.g. NADH, F_{420}, Ferredoxin); **B** – Succinate:quinone oxidoreductase (SDH); **C** – Electron transfer flavoprotein:quinone oxidoreductase (ETF-QO); **D** – Type-2 NADH:quinone oxidoreductase (NDH-2); **E** - Sulfide:quinone oxidoreductase (SQR); **F** – dihidroorotate:quinone oxidoreductase (DHODH); **G** – L-Proline:quinone oxidoreductase (PRODH); **H** – D-amino acid:quinone oxidoreductase (DAADH); **I** – Glycerol-3-phosphate:quinone oxidoreductase (G3PDH-GlpD); **J** – Glycerol-3-phosphate:quinone oxidoreductase (G3PDH-GlpABC); **K** – Malate:quinone oxidoreductase (MQO); **L** – Pyruvate:quinone oxidoreductase (PQO); **M** – Lactate:quinone oxidoreductase (LQO); **N** – Glucose:quinone oxidoreductase (mGDH)/Glycerol:quinone oxidoreductase (GLDH); **O** – Alcohol/Aldehyde:quinone oxidoreductases (ADH/ALDH); **P** – Sorbitol:quinone oxidoreductase (SLDH); **Q** – Gluconate:quinone oxidoreductase (GADH); **R** – Protoporphyrinogen IX:quinone oxidoreductase (PPOR); **S** – Group-1 [NiFe] hydrogenases (exemplifying cytochrome *b* as the membrane subunit); **T** – Formate:quinone oxidoreductase (Fdn-N); **U** – Cytochrome c_3:quinone oxidoreductase (QRC); **V** – Cytochrome:quinone oxidoreductase (Nhc); **W** – Na^+ translocating NADH:quinone oxidoreductase (NQR). P5C ((S)-1-pyrroline-5-carboxylate); G3P (Glycerol-3-phosphate); DHAP (Dihydroxyacetone phosphate); Cyt *c* (Cytochrome *c*); Tpl c_3 (Type I cytochrome c_3) (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively).

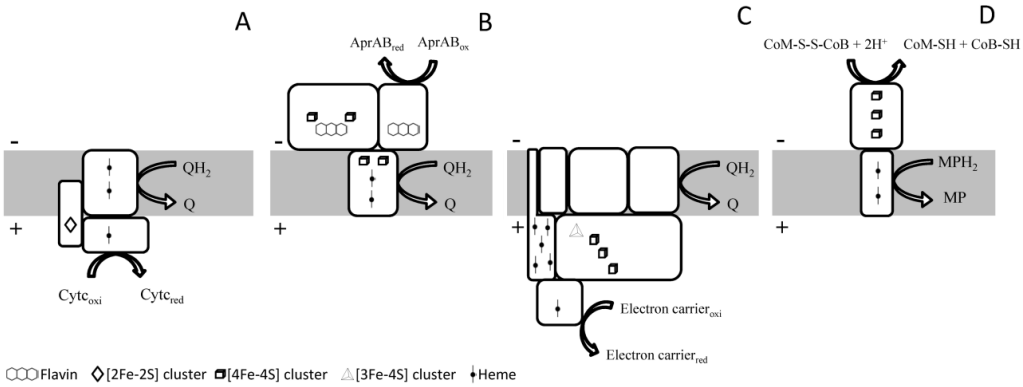


Figure IV.1.3: Quinol oxidases present in respiratory chains. Schematic representation of respiratory proteins or protein complexes involved in quinol oxidation and reduction of soluble electron carriers. Protein representation is proportional to its MM. **A** – Cytochrome bc_1 complex, which is structurally similar to the cytochrome b_6f complexes; **B** – Quinol:electron acceptor oxidoreductase (Qmo); **C** – Alternative Complex III (ACIII); **D** - Methanophenazine:Heterodisulfide oxidoreductase (HDR). Cyt*c* (Cytochrome *c*); AprAB (Adenosine phosphosulfate reductase); CoM-SH (coenzyme M); CoB-SH (coenzyme B); CoM-S–S-CoB (heterodisulfide of coenzyme M and coenzyme B) (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively).

Terminal electron acceptor reductases are the last enzymes present in respiratory chains. In the case of aerobic chains those enzymes are oxygen reductases. Classically only Complex IV (HCOs, Cytochrome *bd* and Alternative oxidase (AOX)) was considered an O_2 reductase, but other enzymes are able to reduce oxygen to water. Prokaryotes may adapt their metabolism to use other final electron acceptors, such as nitrogen, sulfur and organic compounds: nitric oxide reductase (NOR), DMSO reductase, Nitrate reductases, TMAO reductase, Nitrite reductases (Nar/Nap) and Ammonia monooxygenase/ Particulate methane (AMO/pMMO) monooxygenases (Figure IV.1.1 and Figure IV.1.4).

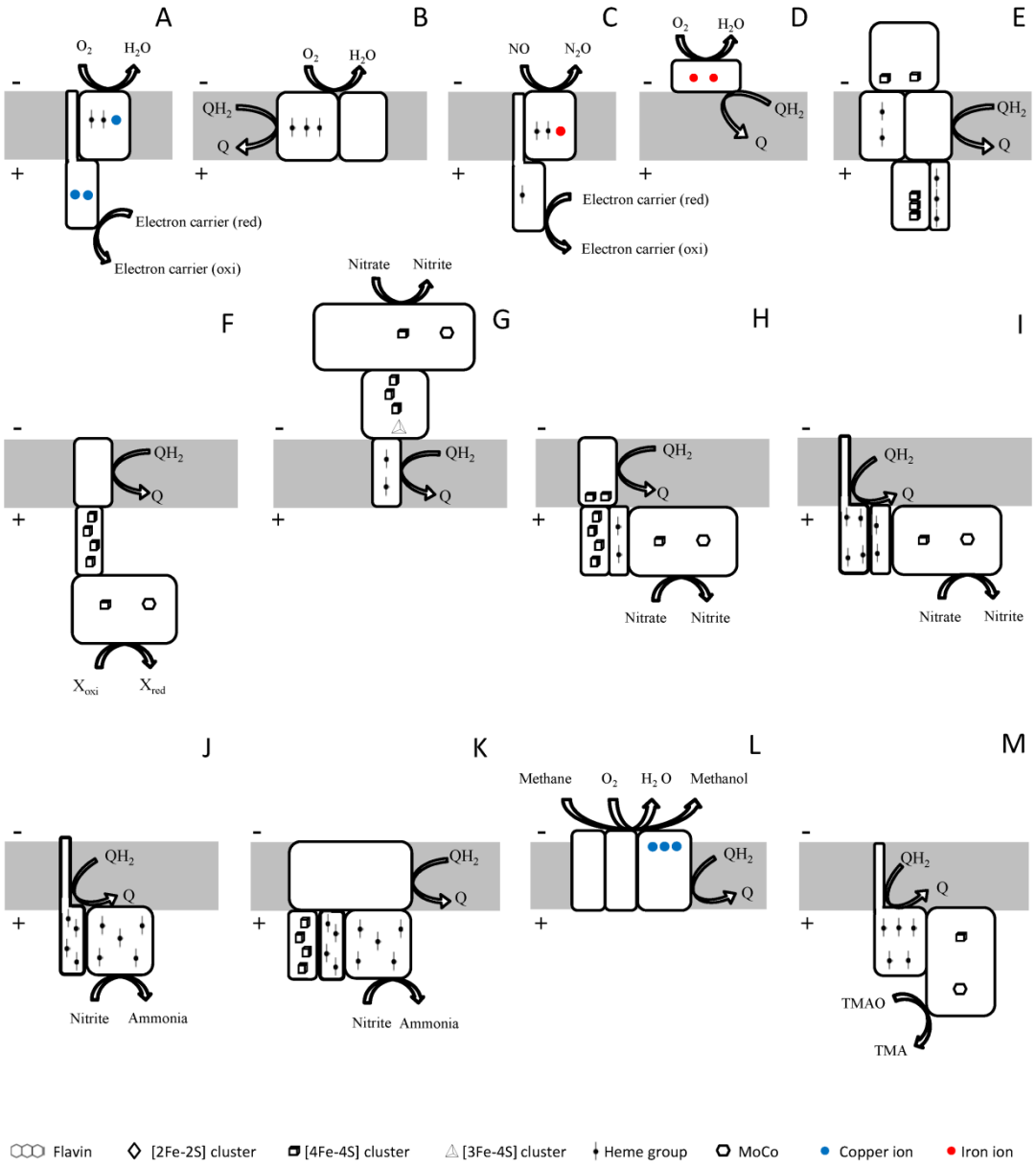


Figure IV.1.4: Final electron acceptor reductases of respiratory chains. Schematic representation of respiratory proteins/complexes involved in reduction of final electron acceptors. In case of homo-oligomers, only monomers are represented. Protein representation is proportional to its MM. **A** – HCO cytochrome oxidase; **B** – cytochrome *bd* quinol oxidase; **C** – cytochrome *c* nitric oxide oxidoreductase (cNOR); **D** – Alternative oxidase (AOX); **E** – Dissimilatory sulfite-reductase complex (Dsr); **F** –

Complex Iron-Sulfur molybdoenzyme (CISM) family (DMSO reductase, Polysulfide reductase, Tetrathionate reductase, Thiosulfate reductase, Sulfur reductase) ; **G** – Quinol:nitrate oxidoreductase (NarGHI); **H** – Periplasmic quinol:nitrate oxidoreductase (NapABGH); **I** – Periplasmic quinol:nitrate oxidoreductase (NapABC); **J** – Cytochrome *c* nitrite reductase (NrfHA); **K** – Cytochrome *c* nitrite reductase (NrfABCD); **L** – Particulate methane monooxygenase (pMMO), representing Ammonia monooxygenase and methane monooxygenase family; **N** – Quinol:TMAO oxidoreductase. Cyt *c* (Cytochrome *c*); DMSO (dimethyl sulfoxide); DMS (Dimethyl sulfide); TMAO (trimethylamine N-oxide); TMA (trimethylamine). X represents different electron acceptors that can be consulted on section I.1 (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively).

Sodium ion-translocating decarboxylases (OAD), Methyltetra-hydromethanopterin: coenzyme M methyltransferase (Mtr) and Membrane-bound pyrophosphatases (mPPase) are ion-translocating oxidoreductases that couple reactions without involving quinones/quinols as substrates to ion translocation across the membrane contributing to the establishment and maintenance of the membrane potential. The bacterial Rnf, which is described to have ferredoxin:NAD^+ oxidoreduction activity in Bacteria and ferredoxin:methanophenazine oxidoreductase activity in Archaea, and members of Complex family, including group-4 membrane-bound [NiFe] hydrogenases (Hyc, Hyf, Ech, mbh and Eha/Ehb) and non-hydrogenases (mbx and Ehr) were also included in this division (Figure 2.2.1, IV.1.1 and IV.1.5).

We also included in our study two enzymes that usually perform their activities by dissipating membrane potential. ATP synthases couple the dissipation of the membrane potential to the synthesis of ATP, being considered the last complex, Complex V, of mitochondrial respiratory chains. The enzyme may operate in the

reverse way and in fact in some organisms this is the only process by which membrane potential may be established. The so-called transhydrogenase dissipates membrane potential to reduce $NADP^+$ by NADH, but may also operate in the reverse way pumping protons (Figure IV.1.1 and Figure IV.1.5).

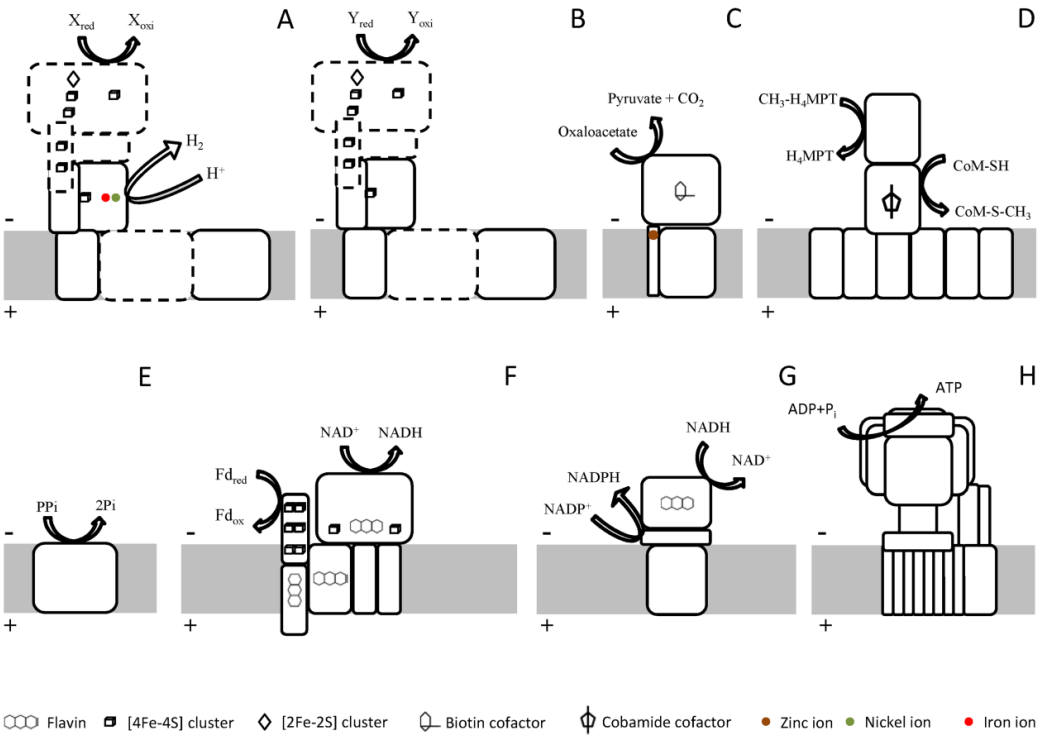


Figure IV.1.5: Ion-Translocating Oxidoreductases and ATP synthases. Schematic representation of membrane proteins or protein complexes involved in ion translocation across the membrane, without involving reaction with quinone/quinol. Protein representation is proportional to its MM. **A** – Complex I family, including group-4 membrane-bound [NiFe] hydrogenases (Hyc, Hyf, Mbh, Ech, Eha and Ehb). Dashed subunits are not present in all enzymes; **B** – Complex I family, including non-hydrogenases (Mbx and Ehr). Dashed subunits are not present in all enzymes; **C** – Oxaloacetate decarboxylase (OAD), representing Na^+- translocating decarboxylases; **D** – Methyltetrahydromethanopterin:coenzyme M methyltransferase (Mtr); **E** – Membrane-bound pyrophosphatase (mPPase); **F** – Rnf complex. **G** – Membrane-bound transhydrogenase (Pnt); **H** – ATP synthase/ATPase. CH_3-H_4MPT (methyltetrahydromethanopterin); H_4MPT (tetrahydromethanopterin); CoM-SH

(coenzyme M); CoM-S-CH_3 (methyl-coenzyme M); PPi (inorganic pyrophosphate); Pi (inorganic phosphate); Fd (Ferredoxin) (+ and – indicate the positive and negative sides of the transmembrane difference in electrochemical potential, respectively).

IV.1.3.1 Membrane-bound Respiratory complexes in Eukarya

The taxonomic profiles, presented in Figure IV.1.1 show all species of animals have genes coding for the classic respiratory complexes, namely Complexes I, II, III and IV and cytochrome *c*; in addition all species also contain genes encoding DHODH and PRODH. Membrane-bound transhydrogenase (Pnt) is present in all animals, with the exception of Placozoans. Furthermore, NDH-2, SQR, G3PDH and ETF-QO are present in between 40 and 100 % of the animal species. Some species also have genes encoding AOX.

The crystallographic structures of the classic mammalian respiratory complexes were already determined including Complex I from *Bos taurus* (PDB code 4UQ8) [11], Complex II from *Sus scrofa* (PDB code 1ZOY) [12], Complex III from *Bos taurus* (PDB code 1BGY) [13] and Complex IV from *B. taurus* (PDB code 1OCC) [14], in addition to that of F_1F_O-ATP synthase from *B. taurus* [15].

Plant membrane-bound respiratory enzymes contrast with those from animals by the presence of NDH-2 and mPPases in all species, as well as of AOX (with the exception of *Micromonas pusilla*), and the complete absence of SQRs. Pnt is only present is some Green and Red algae species. In addition all Plant species contain cytochrome b_6f complexes and type 1 copper proteins as electron carriers, which most probably are present in chloroplasts.

The members of the Fungi phylum Microsporidians (5 species) do not have any component of membrane-bound respiratory chains, what reflects their parasitic nature. The members of the other two phyla from Fungi seem to contain Complexes II, III and IV and also ETF-QO, DHODH and PRODH. Complex I is only present in 20 % and 11 % of the species from the phyla Ascomycetes and Basidiomycetes, respectively, but NDH-2 is observed in all species (excepting *Moniliophthora perniciosa*).

In Protists, the species from the phyla Parasalids and Diplomonads (1 species each) do not have any component of membrane-bound respiratory chain. These organisms are described as parasites and thus most probably obtain energy from their hosts. In the other protist phyla, SDH, NDH-2, G3PDH, ETF-QO, DHODH and PRODH are the main electron delivers to the respiratory chain. In some cases the presence of SQR, LQO and DAADH is also observed. Most members of those phyla also have cytochrome bc_1 complex and the O_2 reductases, HCO and AOX. For example *Plasmodium falciparum* contains a minimal mtDNA and carries out oxidative phosphorylation in the insect vector stages. The reduction of quinone is performed by five dehydrogenases, namely NDH-2, SDH, G3PDH, MQO, and DHOD that do not perform ion translocation. The remaining of the respiratory chain is composed of the canonical Complexes III and IV [16, 17]. In the blood stages of mammalian hosts, mitochondrial enzymes are down-regulated and parasite energy metabolism relies mainly on glycolysis [18, 19]. In *T. brucei*, respiratory Complexes I, III and IV are present and functional in the life cycle stage within the

tsetse fly (vector) but it is absent within the host, where no ATP is generated by oxidative phosphorylation [20-22]. In contrast, a G3PDH and an AOX are present in both stages [23-25].

IV.1.3.2 Membrane-bound Respiratory complexes in Bacteria

Our results highlight the diversity and flexibility of the energetic metabolism of prokaryotes, allowing them to adapt and colonize the most diverse and changing habitats.

Proteobacteria is the largest and phenotypically most diverse group of prokaryotes with sequenced genomes. In addition to mitochondrial classic complexes, proteobacteria have a plethora of enzymes that allow them to use different electron donors and acceptors. Gammaproteobacteria-Enterobacteria species, besides oxidizing NADH (Complex I, 97 % and NDH-2, 91 %) and succinate (SDH, 94 %), use other substrates as electron donors being the most representatives, glycerol-3-Phosphate (G3PDH, 89 %), dihidroorodate (DHODH, 96 %), protoporphyrinogen (PPOR, 92 %), amino acids (PRODH, 79 % and DAADH, 84 %) and D-lactate (LQO 84 %). We observe that almost all (98 %) of the species of Gammaproteobacteria-Enterobacteria bypass cytochrome bc_1 complex (Complex III) and electrons are directly transferred from quinol to the terminal electron acceptor reductases. In fact, only two species (*Citrobacter koseri* and *Citrobacter amalonaticus*) have the genes coding for cytochrome bc_1 complex subunits. Enterobacteria species have genes coding for different oxygen reductases: cytochrome *bd* oxidase (90 %) and HCO (93 %). In additional, several terminal electron acceptors can be used

in replacement of oxygen, such as DMSO (DMSO reductase, 57 %), nitrate (Nar, 63 % and Nap, 47 %) and nitrite (Nrf, 51 %). *E. coli* is a good example of the diversity and flexibility of enterobacterial respiratory chains. *E. coli* is able to use 10 different electron donors by 15 different quinone reductases. Among these are Complex I, NDH-2, formate dehydrogenase, SDH, glucose dehydrogenase, G3PDH (GlpD and GlpABC), DAADH, LQO and group-1 [NiFe] Hydrogenase. *E. coli* can use oxygen, nitrate, nitrite, fumarate, TMAO, DMSO, tetrathionate and thiosulfate as terminal electron acceptors. It has three different quinol:oxygen reductases expressed depending of oxygen availability [26]: bo_3 is expressed under high oxygen concentrations, while cytochromes *bd* are expressed under low oxygen tensions. Both periplasmic (Nap) and respiratory nitrate reductase (Nar) participate in nitrate respiration while a cytochrome *c* nitrite reductase is used for nitrite reduction [27]. Moreover, *E. coli* can synthesize both ubiquinone and menaquinone. In Gammaproteobacteria-Enterobacteria, a difference in the membrane potential may also be created by Rnf (73 %) or by complex I-like complexes (60 %), which includes two members of group-4 [NiFe] Hydrogenase, Hyf and Hyc [8]. The composition of the electron transfer chain of Enterobacteria species is different from that of the Gammaproteobacteria-Other species. In 43 % of these other species, genes coding for Complex I are absent, being its oxidoreductase activity performed by NDH-2 and/or Na^+-NQR. Like Enterobacteria, the other clades of Gammaproteobacteria also use additional substrates as electron donors being the most representatives Glycerol-3-Phosphate (GlpD, 76

%), Dihidroorodate (DHODH 86 %), L-proline (PRODH, 78 %), and other amino acids (DAADH 57 %) and alcohol and aldehydes (alcohol/aldehyde:quinone oxidoreductase, 62%). Electrons from quinol are then transferred to oxygen reductases (89 %) via the cytochrome bc_1 complex (78 %) and soluble electron donors (cytochrome *c* (69 %), HiPIP (14 %) or type 1 copper proteins (48 %)) or directly to quinol:terminal electron acceptors oxidoreductases (*bd* oxidase, 89 %). For example *P. aeruginosa* may use aerobic or anaerobic respiration or fermentation as part of its energetic metabolism [28]. In this bacterium, 17 different dehydrogenases have been identified, including complex I, NDH-2, NQR, formate dehydrogenase and gluconate dehydrogenase. For aerobic respiration, besides cytochrome bc_1 complex and cytochrome *c*, five terminal oxygen reductases have been described: one cytochrome *bd* and four HCOs (three cytochrome *c* oxidases and one quinol oxidase). In anaerobiosis, *P. aeruginosa* is a denitrifying bacterium and enzymes such as nitrate reductase, nitrite reductase, nitric oxide reductase and a nitrous oxide reductase are expressed [29]. In this group of Gammaproteobacteria, 56 % of the species also have Rnf that allows the establishment of the membrane potential.

Species from the Betaproteobacteria clade have, not only the classic components of the mitochondrial electron transfer chain (Complexes I, II, III and IV), but also genes coding for several quinone reductases, such as NDH-2 (76 %), GlpD (78 %), ETF-QO (90 %) and DAADH (84 %). Also, electron transfer between cytochrome bc_1 complex and HCO can be mediated by cytochrome c, HiPIP or type 1

copper proteins. Electrons can be transferred directly from quinol to oxygen via quinol oxidases or to nitrate via nitrate reductase. 82 % and 13 % of the betaproteobacterial species have the genes coding for Pnt and Rnf, respectively.

In comparison with Gamma- and Betaproteobacteria, Epsilonproteobacteria appear to have a less flexible electron transfer chain. Nevertheless, the species analyzed, besides Complexes I and II, have genes coding for MQO (78 %), DHODH (100 %) and group-1 [NiFe] hydrogenenase (95 %). Electron transfer between cytochrome bc_1 complex and HCO is performed by type 1 copper proteins. The presence of genes coding for NapABCGH, NrfAH and TMAO reductase indicates that epsilonproteobacteria are able to use nitrate, nitrite and TMAO as electron acceptors. In the respiratory chain of *Campylobacter jejuni,* a human pathogen, Complex I was shown to oxidize a flavodoxin instead of NADH [30]. *C. jejuni* can also use hydrogen, formate, succinate, malate, lactate and α-ketoglutarate as electron donors [31]. *C. jejuni* is able to perform nitrate, nitrite and TMAO respiration. Nevertheless, this bacterium seems not to be able to grow in strictly anaerobic conditions [32], apparently due to the presence of an oxygen-requiring ribonucleotide reductase [33].

Deltaproteobacteria include a group of aerobic organisms like *Myxobacteria* species, another group defined as strictly anaerobic containing most of known sulfate-reducing bacteria (e.g. *Desulfovibrio* and *Desulfobacter*) and sulfur-reducing bacteria (e.g. *Desulfuromonas* spp.) and a third group of anaerobic organisms (e.g. *Geobacter spp.* ferric iron-reducing). Our results show that the respiratory chains of

Deltaproteobacteria are composed of Complex I (79 %), SDH (91 %), GlpD (72 %) and cytochrome *bd* oxidase (86 %). The metabolism of sulfate- and sulfur-reducing bacteria involves specific enzymes. Therefore, 29 % of the species have QRC, 31 % have Nhc, 34 % have QMO and 40 % contain DSR. These enzymes are, apparently, exclusive of sulfate-reducing bacteria. The quinol:electron carrier oxidoreductase activity can be performed by a bc_1 complex (33 %) or an ACIII (31 %). Additionally, we observed 67 % of the Deltaproteobacteria species have genes coding for mPPases while 55 % have genes coding for complex I family enzymes. Interestingly, although the sulfate- and sulfur-reducing organisms are described as strictly anaerobic, we found genes coding for cytochrome *bd* and type A HCO enzymes, which had already been described, for example, for *Desulfovibrio gigas* [34] and *D. vulgaris*, respectively [35, 36].

Alphaproteobacteria electron transfer chain resembles that of typical mitochondrial chain, composed of Complexes I, II, III and IV and also GlpD, ETF-QO oxidoreductase, PRODH and DAADH. The major differences reside in the presence of genes coding for cytochrome *bd* oxidase (72 %), type 1 copper proteins (39 %) and AMO/pMMO (39 %). An example of an alphaproteobacterial respiratory chain is that of the soil bacterium *P. denitrificans* [37, 38]. In addition to the enzymes homologous to the classic complexes of mitochondria, *P. denitrificans* has a HCO (type C, a cbb_3 enzyme) and can receive electrons from hydrogen via [NiFe] hydrogenase. Anaerobically, *P. denitrificans* can use nitrate (Nap and Nar), nitrite (soluble nitrite reductase), nitric oxide (NOR) and nitrous oxide (soluble N_2O reductase).

Our results indicate that species belonging to Armatimonadetes, Bacteroidetes, Chlamydiae, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Gemmatimonadetes and Planctomycetes phyla have a great variety of quinone reductase enzymes. Nevertheless, electron input into the respiratory chain is performed mainly by Complex I, SDH and NDH-2. In cyanobacteria, Complex I exists in an incomplete form (Complex I-like enzyme), missing the classic electron input module (subunits NuoE, F and G, *E. coli* nomenclature) [39]. Despite the fact that succinate dehydrogenase activity has been identified in cyanobacteria [40], we only found genes coding for peripheral subunits SdhA and SdhB and thereby we did not consider the existence of A-D types SDH in these organisms. GADH, mGDH/GLDH and ADH/ALDH are present in the two full sequenced species of Gemmatimonadetes phylum. Additionally, mPPases appear to play a key role in the metabolism of the organisms belonging to the group of phyla mentioned above, except in Cyanobacteria, Deinococcus-Thermus and Chlamydiae. Pnt was detected in 50 %, 91 % and 53 % of species belonging to Gemmatimonadetes, Cyanobacteria and Deinococcus-Thermus, respectively. As soluble electron carriers, these organisms use copper proteins and/or cytochrome *c*. Despite the great diversity in terms of quinone reductases, members of this group are poorly diverse relatively to terminal electron acceptor reductases, presenting only oxygen reductases, HCO and cytochrome *bd* oxidase. Almost all species have F-ATPase. However, in Deinococcus-Thermus phylum, we only observe A/V-ATPase, as reported before [41, 42]. A minimal

respiratory chain, composed of NQR, SDH and cytochrome *bd* oxidase, appears to be present in the majority of organisms belonging to Chlamydiae, in agreement with the literature [43]. Additionally, we observe the presence of HCOs as terminal electron reductases in some organisms belonging to Chlamydiae, as reported for *Protochlamydia amoebophila* [43].

Interestingly, in the organisms of Bacteroidetes phylum ACIII (69 %) is the only quinol:electron carrier oxidoreductase. The respiratory chain of *R. marinus,* a thermophilic organism belonging to the Bacteroidetes phylum*,* has been extensively studied and was shown to have the classic Complexes I and II [44-47] and also two members of HCO family, caa_3 [48, 49] and ba_3 [50]. Quinol:HiPIP/cytochrome *c* oxidoreductase activity is performed by ACIII [51-54]. The putative respiratory chain of *Porphyromonas gingivalis*, a secondary colonizer of the oral cavity of humans was previously described [55]. Two sodium ion-translocating decarboxylases were identified in this organism using methylmalonyl-CoA and oxaloacetate as substrates. Genes coding NQR and Rnf are observed in the genome of *P. gingivali*. Complex II is encoded by the *frdCAB* operon, indicating that fumarate can be the final electron acceptor. In addition the genome of *P. gingivalis* contains genes coding for a cytochrome *bd* oxidase and for a nitrite reductase (NrfAH) [55]. The species of Chloroflexi phylum have Complex I (100 %), SDH (71 %), G3PDH (71 %), DAADH (7 %), HCOs (71 %) and mPPases (93 %). In addition, approximately half of the species contain genes coding for NDH-2, SQR and ion-translocating decarboxylases. The respiratory

chain of *C. aurantiacus* has been explored and was shown to contain ACIII, which interacts with auracyanin [56-58]. Its menaquinol:fumarate oxidoreductase (mQFR) was exhaustively studied, including the crystallographic structural characterization [59].

Organisms belonging to Acidobacteria, Actinobacteria, Aquificae, Bacilli, Chlorobi, Chrysiogenetes, Deferribacteres, Nitrospirae, Thermodesulfobacteria and Verrucomicrobia also present a great variety of quinone reductase enzymes. The main quinone reductases in these organisms are Complex I, SDH and NDH-2. Additionally, we observed a significant presence of SQR, DAADH, GlpD, PQO and group-1 [NiFe] hydrogenases. A great variety of non-quinone interacting ion-translocating enzymes, such as Complex I family, decarboxylases, mPPases and Rnf, are detected in these phyla, except in Bacilli. Our results indicate that organisms belonging to this group use mainly cytochrome bc_1 complex and ACIII for the oxidation of quinol and reduction of soluble electron carriers. In Chlorobi phylum, Qmo is used in addition to cytochrome bc_1 complex and in Thermodesulfobacteria only Qmo is present, which is in agreement with the fact that these organisms have also sulfur metabolism.

Except in Actinobacteria, Aquificae and Thermodesulfobacteria phyla, soluble electron carriers (type 1 copper proteins and/or cytochrome *c*) are present in all phyla. In Actinobacteria, establishment of a cytochrome bc_1-aa_3 supercomplex has been described, which may explain the absence of soluble cytochromes *c* or other soluble electron carriers [60, 61]. Most phyla have HCO and cytochrome *bd* oxygen reductases, except Thermodesulfobacteria,

which only have cytochrome *bd* oxidase. Interestingly, we found genes coding for HCOs and cytochrome *bd* oxidases in organisms that are described as strictly anaerobes (*Denitrovibrio acetiphilus*, *Calditerrivibrio nitroreducens*, *Desulfurispirillum indicum*, for example) [62-64]. Additionally, we observed genes coding for: terminal electron acceptor reductases related to nitrate metabolism in more than 25 % species of Actinobacteria, Verrucomicrobia, Deferribacteres, Bacilli and Thermodesulfobacteria phyla; terminal electron acceptor reductases related to sulfur and sulfate metabolism in more than 25 % species belonging to Deferribacteres, Chlorobi, Aquificae and Thermodesulfobacteria phyla; and AMO/pMMO monooxygenase in Bacilli and Nitrospirae phyla. In genus *Bacillus* the electron transfer from quinol to oxygen can occur by a diverse set of oxidases as observed in our taxonomic profiles (Figure IV.1.1) and as described in the literature [65, 66]. *Bacillus subtilis* can also grow anaerobically using nitrate as final electron acceptor [67, 68], in agreement with the identification of nitrate reductases. All species of Deferribacteres phylum possess NQR and Rnf, which may be involved in the oxidation of butyrate and in the generation of reduced ferredoxin. Deferribacteres species may also use H^+ as final electron acceptor by [NiFe] hydrogenases [69].

In 50 % of the species from Aquificae phylum, genes coding for NOR are present, allowing the utilization of NO as terminal electron acceptor. For *A. aeolicus*, a hyperthermophilic bacterium, three electron transfer chains have been proposed considering the different electron donors and acceptors. Using the couple H_2/S^0, a

supercomplex composed of a sulfur reductase and a hydrogenase was observed. A respiratory chain composed of SQR, cytochrome bc_1 complex, soluble cytochrome *c*, cytochrome *c* oxidase and a quinol oxidase is expressed when H_2S is used as electron donor and oxygen as terminal electron acceptor. Another chain was proposed for the couple H_2/O_2 and involves a [NiFe] hydrogenase, cytochrome bc_1 complex, a soluble cytochrome *c* and cytochrome *c* oxidase [70]. Members of the Chlorobi phylum, green sulfur bacteria (GSB), are photoautotrophic bacteria, all containing Complexes I and II, as well as SQRs and mPPases. Quinone reduction may also be performed by ETF-QO (present in 58 % of the species) and group-1 [Ni-Fe] hydrogenase (58 %). Besides SQR, other proteins of sulfur metabolism have been detected, such as Dsr and polysulfide reductase, which are present in 25 % and 100 % of the species, respectively. We also observed the presence of HCOs and cytochrome *bd* oxidase in 50 and 83 % of the species of Chlorobi phylum, indicating that these organisms may also use O_2 as final electron acceptor. Most species (83 %) have a cytochrome bc_1 complex, while 25 % have a Qmo complex. Chlorobi may also establish membrane potential by Na^+-dependent decarboxylases (67 % of the species), Rnf (58 %) and NQR (8 %). Species from the phylum Thermodesulfobacteria have been described as sulfate-reducing organisms. In fact, our results indicate the presence of enzymes that are involved in that metabolism. We observe the presence of Complex I, Complex I family enzymes and group-1 [NiFe] hydrogenases as electron input enzymes, Qmo as quinol:soluble electron carrier oxidoreductase and cytochrome *bd*

oxidase, DMSO reductase, nitrite reductase and polysulfide reductases as terminal electron acceptor reductases. Although these organisms do not perform nitrite ammonification, some species possess NrfHA complex.

In Clostridia, known as acetogenic bacteria, we identified the genes coding for complex I family enzymes (42 %), Na^+-translocating decarboxylase complex (48 %), mPPase (70 %) and Rnf complex (49 %), which indicates most organisms rely on the membrane potential established by redox-driven ion translocation, not dependent on quinone/quinol. Moreover, we also found genes coding for ATP synthases. *A. woodii* seems to establish a Na^+ dependent membrane potential, probably by Rnf complex and mPPase, which is converted into ATP by a Na^+-specific ATPase [71, 72]. *Clostridium ljungdahlii* [73, 74] conserves energy exclusively through a H^+ dependent membrane potential, possibly generated by Rnf complex [75], coupled to ATP synthesis by a H^+-translocating ATP synthase [73]. In *Moorella thermoacetica*, besides complex I family enzymes and mPPase, also two cytochromes *b* and menaquinone-7 were described as being involved in the generation of a H^+ dependent membrane potential [76] and a subunit of cytochrome *bd* oxidase was detected, with unclear function [76].

The closely related Thermotogae, Synergistetes and Dictyoglomi phyla are composed of anaerobic bacteria [77] and are similar, in terms of membrane-bound respiratory complexes, to Caldiserica, Fusobacteria and Elusimicrobia phyla. Our results show that the majority of these species have genes coding for subunits or

proteins involved in the establishment of transmembrane potential, such as Complex I family enzymes, Rnf, mPPase and Na^+-translocating decarboxylase complex, and ATP synthases. Membrane-bound PPase from the hyperthermophilic bacterium *T. maritima* was biochemically characterized [78]. Glutaconyl-CoA decarboxylase from the anaerobic bacterium *Fusobacterium nucleatum* performs glutamate decarboxylation and generates a Na^+ dependent membrane potential, which can be coupled to ATP synthesis, via a F-type ATP synthase [79, 80].

Spirochetes, are widespread in aquatic environments and in animals and can cause several diseases, such as syphilis (*Treponema pallidum)*, leptospirosis (*Leptospira* sp.) and Lyme disease (*Borrelia burgdorferi*). In anaerobic species, such as *B. burgdorferi* and *T. pallidum*, our results indicate the presence of genes coding for proteins involved in the establishment of the membrane potential, such as Rnf complex, mPPase and Na^+-translocating decarboxylase complex, GlpD and ATP synthase. *B. burgdorferi* lacks a respiratory chain and ATP production is accomplished by fermentation. The membrane potential is established by the reverse reaction of the V-type ATPase [81, 82]. In species from *Leptospira,* which includes aerobic organisms, we detected the presence of NADH dehydrogenases, GlpD, ACIII and HCO.

In organisms belonging to Tenericutes phylum, only genes coding for ATPase are identified. *Mycoplasma pneumoniae*, a human pathogen with a reduced genome and parasitic life [83], lacks a functional respiratory chain and relies on fermentation for ATP

generation. In this bacterium, F-ATPase creates a membrane potential in order to maintain cell homeostasis and to allow nutrient import [84].

IV.1.3.3 Membrane-bound Respiratory complexes in Archaea

Archaea in KEGG database are divided into six main phyla, three of which are composed of only one species, Nanoarchaeota, Korarchaeota and Lokiarchaeota. According to our results, none of these three phyla have membrane soluble quinones. Nonetheless, these are archaeal phyla and thus we cannot exclude the existence of other quinone systems. For example the benzothiophenquinone, which biosynthesis pathway is not known yet, is present in some organisms from Crenarchaeota. Nanoarchaeota respiratory chain seems to be highly limited, considering that only A/V-ATPase and SQR are present.

Euryarchaeota phylum contains the majority of the sequenced archaeal genomes (118 out of 173 species) and includes 3 main groups: Methanogenic Archaea, Halobacteria and Thermococcus, covering 105 from the 118 species. Methanogenic Archaea produce methane in their anaerobic respiratory metabolism [85]. This group includes some archaeal genera like *Methanocaldococcus*, *Methanococcus*, *Methanosarcina*, *Methanosaeta* and *Methanobacterium*. All 55 species of methanogenic archaea have A/V-ATP synthase and MTR and most of them (87 %) contain Complex I-like enzymes. Our results also show that three other ion translocating systems are found in these organisms: Rnf (12 out of 55 species),

mPPases (30 out of 55 species) and group-1 [NiFe]-hydrogenases (restricted to the 12 species, 100 %, of the *Methanosarcina* family) [86]. Presence of methanophenazine in the membranes of these organisms was also described [87]. We have identified three protein complexes that work as electron input systems: LQO (49 %), Complex I-like enzymes (87 %) and Rnf (22 %). We observed the presence of HDR (38 %) [88], and of cytochrome *bd* oxidase (13 %), although these methanogenic archaea are described as anaerobes. In *M. mazei*, ATP synthase, MTR complex, [NiFe]-hydrogenases and Rnf were identified [89], in agreement with our results. The second considered Euryarchaeota group is Halobacteria, which are characterized by having a natural environment with high salt concentration [90]. We observed that all the species of this group have A/V-ATP synthase, as well as the presence of HCO in 28 species, cytochrome bc_1 complex in 17 species and cytochrome *bd* oxidase in 20 out of the 29 species. We identified SDH (100 %), NDH-2 (97 %), MQO (48 %), ETF-QO (76 %), PRODH (83 %), LQO (100 %), DHODH (100 %), SQR (55 %) and GlpABC (76 %). These results reflect the ability of these organisms to feed electrons into the quinone pool from different metabolic origins. Halobacteria seem to have a well defined chain for quinol oxidation. They have the ability to directly transfer electrons from quinol to terminal electron acceptors by HCO and/or NOR (52 % of the species) or they can oxidize quinol with reduction of soluble electron carriers (HiPIP or type 1 cooper proteins, present in 12 and 28 species respectively), as supported by the presence of cytochrome bc_1 complex. HCO and NOR may also perform the oxidation of the same

soluble electron carriers and concomitant final electron acceptor reduction (oxygen and nitric-oxide, respectively). Anaerobic respiration systems have also been described, including Nitrate reductases [91] and TMAO/DMSO reductases but our analysis did not identified the corresponding genes in any Halobacteria. Together, these results seem to suggest a high versatility in the electron input proteins, and the ability of these Archaea to perform both aerobic and anaerobic respiration, as it has been reported [92, 93]. If we compare the obtained results with the proposed respiratory chain of *Natronomonas pharaonis* (aerobic Halobacterium), we concluded that all proteins described [94], which include SDH, cytochrome *c*, cytochrome bc_1 complex and cytochrome ba_3 oxidase, were identified in our study. The third considered group is *Thermococcus*, which are characterized by their prevalence in high temperature environments [95]. According to our results, in addition to A/V-ATPase, that is ubiquitous in this group, *Thermococcus* are able to generate membrane potential through Complex I-like enzymes (19 out of 21 species) and a Na^+-translocating decarboxylase (19 out of the 21 species). Other than these protein complexes, the presence of [NiFe]-hydrogenases has been described in the literature [96]. We do not identified any membrane-bound terminal electron acceptor reductase in this group of organisms. Nevertheless, some terminal oxidases have been described in the literature [97].

Crenarcheota is the second largest Archaea phylum in KEGG database, containing 40 species with sequenced genomes, which can also be divided into three main groups: Desulfurococcus, Sulfolobales

and Thermoprotei (with 13, 8 and 17 species, respectively). In contrast to what is observed for Euryarchaeota, the metabolism of Crenarcheota does not seem to differ substantially. The presence of membrane soluble quinones was observed (Figure IV.1.1) [98-100] , as well as of A/V-ATP synthases. In Crenarcheota we identified three other ion translocating systems: Complex I-like enzymes (73 %, mainly in *Desulfurococcus* and Sulfolobales), HCO (38 %, in *Sulfolobales* and *Thermoprotei*) [101] and cytochrome bc_1 complex (48 %) [102]. Our results show cytochrome *c* and/or a type 1 cooper protein are present in 8 and 33 % of the species, respectively; the majority of the species which contains both homologous cytochrome bc_1 complex and a type 1 copper protein are included in *Sulfolobales* and *Thermoprotei* groups. We also identified the presence of SDH (58 %, mainly in *Thermoprotei*) [103], NDH-2 (35 %, mainly in *Sulfolobales*) [104], SQR (73 %) [105], GlpD (30 %), ETF-QO (63 %) and LQO (78 %). NOR is present in 30 % of the species (mainly in *Thermoprotei*). Specifically in the *Sulfolobales* group, the presence of NDH-2, SDH, homologous cytochrome bc_1-like complex and variable terminal oxidases has been previously shown [106, 107], which is in clear agreement with our results.

IV.1.4 Conclusion

We explored the diversity of membrane respiratory chains and the presence of the different enzyme complexes in the several phyla of life. It was our aim to show the large diversity of respiratory

complexes in the three domains of life, bring focus to less discussed enzymes and highlight the role of the membrane potential, i.e. chemiosmotic coupling, to all forms of life. We investigated more than 50 proteins or protein complexes whose coding genes are present in fully sequenced genomes deposited at KEGG database. As all bioinformatic approaches our procedure is limited by the sampling available (KEGG database), the methodology (BLASTp) and the parameters (E-value < 0.01) used to retrieve the sequences. When needed, we further refined sequences analyses using additional considerations, including sequence length, gene clustering, gene cluster organization, sequence alignment, as well as neighbor-joining clustering.

Although the presence of false positives or non detection of some genes cannot be totally excluded, the combined obtained profiles are illustrative of the distribution and diversity of respiratory complexes in the three domains of life.

The mitochondria respiratory chain may include more enzymes than the classic Complexes I, II, III and IV. Genes coding for ETF-QO, NDH-2, DHODH, PRODH, DAADH and G3PDH were observed in all vertebrate species.

Almost all eukaryotic phyla contain organisms with genes coding for oxygen reductases. The exceptions are three phyla, one from Fungi, Microsporidians (5 species) and two from Protists, Parasalids and Diplomonads (one species each), which do not contain any membrane-bound respiratory enzyme. The members of these phyla are described as parasites.

Proteobacteria are the most diverse organisms in terms of the energetic metabolism based on the analyses of the membrane-bound respiratory enzymes. They present genes coding for a variety of enzymes able to use different terminal electron acceptors.

Oxygen seems to be the preferential electron acceptor of respiratory chains. Importantly, all phyla not presenting O_2 reductases also apparently do not have any other type of quinol:final electron acceptor oxidoreductases.

Interestingly, genes coding for enzymes involved in quinone biosynthetic pathways seem to be not present in 14 out of 45 prokaryotic phyla. Accordingly, these phyla seem not to contain quinone/quinol interacting complexes. With the exception of the species from Tenericutes (65) and Nanoarchaeota (1) phyla, the members of the other phyla are able to establish membrane potential through the action of redox driven ion-translocation without the intervention of quinone/quinols. This allows the respective organisms to use the membrane potential for ATP synthesis, solute transport and motility. In this context we propose enzymes that perform redox or decarboxylation driven ion translocation, such as oxaloacetate decarboxylase, Rnf (that does not interact with quinone) and complex I family (e.g. mbx from *T. maritima*) to be considered respiratory complexes.

Excluding totally parasitic phyla, all phyla have at least one gene coding for a redox driven ion pump (Complex I, HCO, decarboxylases, mPPase, Rnf) and/or ATP synthases. Our data show the establishment of membrane potential is vital to life.

IV.1.5 Acknowledgments

FC, PJC, AMD, FVS and PNR are recipients of fellowships by Fundação para a Ciência e a Tecnologia (SFRH/BD/104481/2014, SFRH/BD/97730/2013, SFRH/BPD/78075/2011, PD/BD/113985/2015 and SFRH/BPD/71022/2010). The work was funded by Fundação para a Ciência e a Tecnologia (PTDC/BBB-BQB/2294/2012 to M.M.P). ITQB is supported by Fundação para a Ciência e a Tecnologia through R&D Unit, UID/CBQ/04612/2013.

IV.1.6 References

1. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. & Yamanishi, Y. (2008) KEGG for linking genomes to life and the environment, *Nucleic acids research.* 36, D480-4.

2. Kanehisa, M. & Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes, *Nucleic acids research.* 28, 27-30.

3. Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. & Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG, *Nucleic acids research.* 34, D354-7.

4. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. (2009) BLAST+: architecture and applications, *BMC bioinformatics.* 10, 421.

5. Pires, R. H., Lourenco, A. I., Morais, F., Teixeira, M., Xavier, A. V., Saraiva, L. M. & Pereira, I. A. (2003) A novel membrane-bound respiratory complex from *Desulfovibrio desulfuricans* ATCC 27774, *Biochim Biophys Acta.* 1605, 67-82.

6. Saraiva, L. M., da Costa, P. N., Conte, C., Xavier, A. V. & LeGall, J. (2001) In the facultative sulphate/nitrate reducer *Desulfovibrio desulfuricans* ATCC 27774, the nine-haem cytochrome *c* is part of a membrane-bound redox complex mainly expressed in sulphate-grown cells, *Biochim Biophys Acta.* 1520, 63-70.

7. Pires, R. H., Venceslau, S. S., Morais, F., Teixeira, M., Xavier, A. V. & Pereira, I. A. (2006) Characterization of the *Desulfovibrio desulfuricans* ATCC 27774 DsrMKJOP complex--a membrane-bound redox complex involved in the sulfate respiratory pathway, *Biochemistry.* 45, 249-62.

8. Marreiros, B. C., Batista, A. P., Duarte, A. M. & Pereira, M. M. (2013) A missing link between Complex I and group 4 membrane-bound [NiFe] hydrogenases, *Biochim Biophys Acta.* 1827, 198-209.

9. Pei, J., Kim, B. H. & Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments, *Nucleic acids research.* 36, 2295-300.

10. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics.* 23, 2947-8.

11. Vinothkumar, K. R., Zhu, J. & Hirst, J. (2014) Architecture of mammalian respiratory complex I, *Nature.* 515, 80-4.

12. Sun, F., Huo, X., Zhai, Y., Wang, A., Xu, J., Su, D., Bartlam, M. & Rao, Z. (2005) Crystal structure of mitochondrial respiratory membrane protein complex II, *Cell.* 121, 1043-57.

13. Iwata, S., Lee, J. W., Okada, K., Lee, J. K., Iwata, M., Rasmussen, B., Link, T. A., Ramaswamy, S. & Jap, B. K. (1998) Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex, *Science.* 281, 64-71.

14. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. & Yoshikawa, S. (1996) The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8 A, *Science.* 272, 1136-44.

15. Walker, J. E. (2013) The ATP synthase: the understood, the uncertain and the unknown, *Biochemical Society transactions.* 41, 1-16.

16. Nixon, G. L., Pidathala, C., Shone, A. E., Antoine, T., Fisher, N., O'Neill, P. M., Ward, S. A. & Biagini, G. A. (2013) Targeting the mitochondrial electron transport chain of *Plasmodium falciparum*: new strategies towards the development of improved antimalarials for the elimination era, *Future medicinal chemistry.* 5, 1573-91.

17. Painter, H. J., Morrisey, J. M., Mather, M. W. & Vaidya, A. B. (2007) Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*, *Nature.* 446, 88-91.

18. Mogi, T. & Kita, K. (2010) Diversity in mitochondrial metabolic pathways in parasitic protists Plasmodium and Cryptosporidium, *Parasitology international.* 59, 305-12.

19. Krungkrai, J., Burat, D., Kudan, S., Krungkrai, S. & Prapunwattana, P. (1999) Mitochondrial oxygen consumption in asexual and sexual blood stages of the human malarial parasite, *Plasmodium falciparum*, *The Southeast Asian journal of tropical medicine and public health.* 30, 636-42.

20. Bringaud, F., Riviere, L. & Coustou, V. (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources, *Molecular and biochemical parasitology.* 149, 1-9.

21. Opperdoes, F. R. & Michels, P. A. (2008) Complex I of *Trypanosomatidae*: does it exist?, *Trends in parasitology.* 24, 310-7.

22. Tielens, A. G. & van Hellemond, J. J. (2009) Surprising variety in energy metabolism within *Trypanosomatidae*, *Trends in parasitology.* 25, 482-90.

23. Stebeck, C. E., Frevert, U., Mommsen, T. P., Vassella, E., Roditi, I. & Pearson, T. W. (1996) Molecular characterization of glycosomal NAD^+-dependent glycerol 3-phosphate dehydrogenase from *Trypanosoma brucei* rhodesiense, *Molecular and biochemical parasitology.* 76, 145-58.

24. Clarkson, A. B., Jr., Bienen, E. J., Pollakis, G. & Grady, R. W. (1989) Respiration of bloodstream forms of the parasite *Trypanosoma brucei* is dependent on a plant-like alternative oxidase, *J Biol Chem.* 264, 17770-6.

25. Muller, M., Mentel, M., van Hellemond, J. J., Henze, K., Woehle, C., Gould, S. B., Yu, R. Y., van der Giezen, M., Tielens, A. G. & Martin, W. F. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes, *Microbiology and molecular biology reviews : MMBR.* 76, 444-95.

26. Borisov, V. B. & Verkhovsky, M. I. (2009) Oxygen as Acceptor, *EcoSal Plus.* 3.

27. Cole, J. A. & Richardson, D. J. (2008) Respiration of Nitrate and Nitrite, *EcoSal Plus.* 3.

28. Williams, H. D., Zlosnik, J. E. & Ryall, B. (2007) Oxygen, cyanide and energy generation in the cystic fibrosis pathogen *Pseudomonas aeruginosa*, *Advances in microbial physiology.* 52, 1-71.

29. Arai, H. (2011) Regulation and Function of Versatile Aerobic and Anaerobic Respiratory Metabolism in *Pseudomonas aeruginosa*, *Frontiers in microbiology.* 2, 103.

30. Westfall, H. N., Rollins, D. M. & Weiss, E. (1986) Substrate utilization by *Campylobacter jejuni* and *Campylobacter coli*, *Applied and environmental microbiology.* 52, 700-5.

31. Weerakoon, D. R., Borden, N. J., Goodson, C. M., Grimes, J. & Olson, J. W. (2009) The role of respiratory donor enzymes in *Campylobacter jejuni* host colonization and physiology, *Microbial pathogenesis.* 47, 8-15.

32. Sellars, M. J., Hall, S. J. & Kelly, D. J. (2002) Growth of *Campylobacter jejuni* supported by respiration of fumarate, nitrate, nitrite, trimethylamine-N-oxide, or dimethyl sulfoxide requires oxygen, *Journal of bacteriology.* 184, 4187-96.

33. Pittman, M. S. & Kelly, D. J. (2005) Electron transport through nitrate and nitrite reductases in *Campylobacter jejuni*, *Biochemical Society transactions.* 33, 190-2.

34. Lemos, R. S., Gomes, C. M., Santana, M., LeGall, J., Xavier, A. V. & Teixeira, M. (2001) The 'strict' anaerobe *Desulfovibrio gigas* contains a membrane-bound oxygen-reducing respiratory chain, *Febs Lett.* 496, 40-3.

35. Lobo, S. A., Almeida, C. C., Carita, J. N., Teixeira, M. & Saraiva, L. M. (2008) The haem-copper oxygen reductase of *Desulfovibrio vulgaris* contains a dihaem cytochrome *c* in subunit II, *Biochim Biophys Acta.* 1777, 1528-34.

36. Lobo, S. A., Melo, A. M., Carita, J. N., Teixeira, M. & Saraiva, L. M. (2007) The anaerobe *Desulfovibrio desulfuricans* ATCC 27774 grows at nearly atmospheric oxygen levels, *Febs Lett.* 581, 433-6.

37. Nicholls, D. G. & Ferguson, S. J. (2013) 5 - Respiratory chains in *Bioenergetics (Fourth Edition)* (Ferguson, D. G. N. J., ed) pp. 91-157, Academic Press, London.

38. Van Spanning, R. J., de Boer, A. P., Reijnders, W. N., De Gier, J. W., Delorme, C. O., Stouthamer, A. H., Westerhoff, H. V., Harms, N. & van der Oost, J. (1995)

Regulation of oxidative phosphorylation: the flexible respiratory network of *Paracoccus denitrificans*, *Journal of bioenergetics and biomembranes.* 27, 499-512.

39. Battchikova, N., Eisenhut, M. & Aro, E. M. (2011) Cyanobacterial NDH-1 complexes: novel insights and remaining puzzles, *Biochim Biophys Acta.* 1807, 935-44.

40. Cooley, J. W., Howitt, C. A. & Vermaas, W. F. (2000) Succinate:quinol oxidoreductases in the cyanobacterium *Synechocystis sp*. strain PCC 6803: presence and function in metabolism and electron transport, *Journal of bacteriology.* 182, 714-22.

41. Bernal, R. A. & Stock, D. (2004) Three-dimensional structure of the intact *Thermus thermophilus* H^+-ATPase/synthase by electron microscopy, *Structure.* 12, 1789-98.

42. Lapierre, P., Shial, R. & Gogarten, J. P. (2006) Distribution of F- and A/V-type ATPases in *Thermus scotoductus* and other closely related species, *Systematic and applied microbiology.* 29, 15-23.

43. Bavoil, P. M. & Wyrick, P. B. (2006) *Chlamydia: genomics and pathogenenesis*, Horizon Bioscience, Baltimore.

44. Fernandes, A. S., Sousa, F. L., Teixeira, M. & Pereira, M. M. (2006) Electron paramagnetic resonance studies of the iron-sulfur centers from complex I of *Rhodothermus marinus*, *Biochemistry.* 45, 1002-8.

45. Fernandes, A. S., Pereira, M. M. & Teixeira, M. (2002) Purification and characterization of the complex I from the respiratory chain of *Rhodothermus marinus*, *Journal of bioenergetics and biomembranes.* 34, 413-21.

46. Fernandes, A. S., Konstantinov, A. A., Teixeira, M. & Pereira, M. M. (2005) Quinone reduction by *Rhodothermus marinus* succinate:menaquinone oxidoreductase is not stimulated by the membrane potential, *Biochem Biophys Res Commun.* 330, 565-70.

47. Fernandes, A. S., Pereira, M. M. & Teixeira, M. (2001) The succinate dehydrogenase from the thermohalophilic bacterium *Rhodothermus marinus*: redox-Bohr effect on heme bL, *Journal of bioenergetics and biomembranes.* 33, 343-52.

48. Pereira, M. M., Santana, M., Soares, C. M., Mendes, J., Carita, J. N., Fernandes, A. S., Saraste, M., Carrondo, M. A. & Teixeira, M. (1999) The caa_3 terminal oxidase of the thermohalophilic bacterium *Rhodothermus marinus*: a HiPIP:oxygen oxidoreductase lacking the key glutamate of the D-channel, *Biochim Biophys Acta.* 1413, 1-13.

49. Pereira, M. M., Sousa, F. L., Teixeira, M., Nyquist, R. M. & Heberle, J. (2006) A tyrosine residue deprotonates during oxygen reduction by the caa_3 reductase from *Rhodothermus marinus*, *Febs Lett.* 580, 1350-4.

50. Verissimo, A. F., Pereira, M. M., Melo, A. M., Hreggvidsson, G. O., Kristjansson, J. K. & Teixeira, M. (2007) A ba_3 oxygen reductase from the thermohalophilic bacterium *Rhodothermus marinus*, *Fems Microbiol Lett.* 269, 41-7.

51. Pereira, M. M., Refojo, P. N., Hreggvidsson, G. O., Hjorleifsdottir, S. & Teixeira, M. (2007) The alternative complex III from *Rhodothermus marinus* - a prototype of a new family of quinol:electron acceptor oxidoreductases, *Febs Lett.* 581, 4831-5.

52. Pereira, M. M., Carita, J. N. & Teixeira, M. (1999) Membrane-bound electron transfer chain of the thermohalophilic bacterium *Rhodothermus marinus*: a novel multihemic cytochrome *bc*, a new complex III, *Biochemistry.* 38, 1268-75.

53. Pereira, M. M., Carita, J. N. & Teixeira, M. (1999) Membrane-bound electron transfer chain of the thermohalophilic bacterium *Rhodothermus marinus*: characterization of the iron-sulfur centers from the dehydrogenases and investigation of the high-potential iron-sulfur protein function by in vitro reconstitution of the respiratory chain, *Biochemistry.* 38, 1276-83.

54. Refojo, P. N., Teixeira, M. & Pereira, M. M. (2010) The alternative Complex III of *Rhodothermus marinus* and its structural and functional association with caa_3 oxygen reductase, *Biochim Biophys Acta.* 1797, 1477-82.

55. Meuric, V., Rouillon, A., Chandad, F. & Bonnaure-Mallet, M. (2010) Putative respiratory chain of *Porphyromonas gingivalis*, *Future Microbiol.* 5, 717-34.

56. Gao, X., Xin, Y., Bell, P. D., Wen, J. & Blankenship, R. E. (2010) Structural analysis of alternative complex III in the photosynthetic electron transfer chain of *Chloroflexus aurantiacus*, *Biochemistry.* 49, 6670-9.

57. Yanyushin, M. F., del Rosario, M. C., Brune, D. C. & Blankenship, R. E. (2005) New class of bacterial membrane oxidoreductases, *Biochemistry.* 44, 10037-45.

58. Gao, X., Xin, Y. & Blankenship, R. E. (2009) Enzymatic activity of the alternative complex III as a menaquinol:auracyanin oxidoreductase in the electron transfer chain of *Chloroflexus aurantiacus*, *Febs Lett.* 583, 3275-9.

59. Xin, Y., Lu, Y. K., Fromme, R., Fromme, P. & Blankenship, R. E. (2009) Purification, characterization and crystallization of menaquinol:fumarate oxidoreductase from the green filamentous photosynthetic bacterium *Chloroflexus aurantiacus*, *Biochim Biophys Acta.* 1787, 86-96.

60. Niebisch, A. & Bott, M. (2003) Purification of a cytochrome *bc*-aa_3 supercomplex with quinol oxidase activity from *Corynebacterium glutamicum*. Identification of a fourth subunity of cytochrome aa_3 oxidase and mutational analysis of diheme cytochrome c_1, *J Biol Chem.* 278, 4339-46.

61. Megehee, J. A., Hosler, J. P. & Lundrigan, M. D. (2006) Evidence for a cytochrome *bcc*-aa_3 interaction in the respiratory chain of *Mycobacterium smegmatis*, *Microbiology.* 152, 823-9.

62. Rauschenbach, I., Narasingarao, P. & Haggblom, M. M. (2011) *Desulfurispirillum indicum* sp. nov., a selenate- and selenite-respiring bacterium isolated from an estuarine canal, *International journal of systematic and evolutionary microbiology.* 61, 654-8.

63. Iino, T., Nakagawa, T., Mori, K., Harayama, S. & Suzuki, K. (2008) *Calditerrivibrio nitroreducens* gen. nov., sp. nov., a thermophilic, nitrate-reducing bacterium isolated from a terrestrial hot spring in Japan, *International journal of systematic and evolutionary microbiology.* 58, 1675-9.

64. Myhr, S. & Torsvik, T. (2000) *Denitrovibrio acetiphilus*, a novel genus and species of dissimilatory nitrate-reducing bacterium isolated from an oil reservoir model column, *International journal of systematic and evolutionary microbiology.* 50 Pt 4, 1611-9.

65. Sousa, P. M., Videira, M. A., Santos, F. A., Hood, B. L., Conrads, T. P. & Melo, A. M. (2013) The *bc*:*caa*$_3$ supercomplexes from the Gram positive bacterium *Bacillus*

subtilis respiratory chain: a megacomplex organization?, *Archives of biochemistry and biophysics.* 537, 153-60.

66. Garcia Montes de Oca, L. Y., Chagolla-Lopez, A., Gonzalez de la Vara, L., Cabellos-Avelar, T., Gomez-Lojero, C. & Gutierrez Cirlos, E. B. (2012) The composition of the *Bacillus subtilis* aerobic respiratory chain supercomplexes, *Journal of bioenergetics and biomembranes.* 44, 473-86.

67. Sun, G., Sharkova, E., Chesnut, R., Birkey, S., Duggan, M. F., Sorokin, A., Pujic, P., Ehrlich, S. D. & Hulett, F. M. (1996) Regulators of aerobic and anaerobic respiration in *Bacillus subtilis*, *Journal of bacteriology.* 178, 1374-85.

68. Nakano, M. M. & Zuber, P. (1998) Anaerobic growth of a "strict aerobe" (*Bacillus subtilis*), *Annual review of microbiology.* 52, 165-90.

69. Takaki, Y., Shimamura, S., Nakagawa, S., Fukuhara, Y., Horikawa, H., Ankai, A., Harada, T., Hosoyama, A., Oguchi, A., Fukui, S., Fujita, N., Takami, H. & Takai, K. (2010) Bacterial lifestyle in a deep-sea hydrothermal vent chimney revealed by the genome sequence of the thermophilic bacterium *Deferribacter desulfuricans* SSM1, *DNA research : an international journal for rapid publication of reports on genes and genomes.* 17, 123-37.

70. Guiral, M., Prunetti, L., Lignon, S., Lebrun, R., Moinier, D. & Giudici-Orticonit, M. T. (2009) New insights into the respiratory chains of the chemolithoautotrophic and hyperthermophilic bacterium *Aquifex aeolicus*, *Journal of proteome research.* 8, 1717-30.

71. Aufurth, S., Schagger, H. & Muller, V. (2000) Identification of subunits a, b, and c1 from *Acetobacterium woodii* Na+-F1F0-ATPase. Subunits c1, c2, AND c3 constitute a mixed c-oligomer, *J Biol Chem.* 275, 33297-301.

72. Muller, V., Aufurth, S. & Rahlfs, S. (2001) The Na^+ cycle in *Acetobacterium woodii:* identification and characterization of a Na^+ translocating F(1)F(0)-ATPase with a mixed oligomer of 8 and 16 kDa proteolipids, *Biochim Biophys Acta.* 1505, 108-20.

73. Hugenholtz, J. & Ljungdahl, L. G. (1989) Electron transport and electrochemical proton gradient in membrane vesicles of *Clostridium thermoautotrophicum*, *Journal of bacteriology.* 171, 2873-5.

74. Meier, T., Ferguson, S. A., Cook, G. M., Dimroth, P. & Vonck, J. (2006) Structural investigations of the membrane-embedded rotor ring of the F-ATPase from *Clostridium paradoxum*, *Journal of bacteriology.* 188, 7759-64.

75. Kopke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., Ehrenreich, A., Liebl, W., Gottschalk, G. & Durre, P. (2010) *Clostridium ljungdahlii* represents a microbial production platform based on syngas, *Proceedings of the National Academy of Sciences of the United States of America.* 107, 13087-92.

76. Das, A., Silaghi-Dumitrescu, R., Ljungdahl, L. G. & Kurtz, D. M., Jr. (2005) Cytochrome *bd* oxidase, oxidative stress, and dioxygen tolerance of the strictly anaerobic bacterium *Moorella thermoacetica*, *Journal of bacteriology.* 187, 2020-9.

77. Nishida, H., Beppu, T. & Ueda, K. (2011) Whole-genome comparison clarifies close phylogenetic relationships between the phyla *Dictyoglomi* and *Thermotogae*, *Genomics.* 98, 370-5.

78. Belogurov, G. A., Malinen, A. M., Turkina, M. V., Jalonen, U., Rytkonen, K., Baykov, A. A. & Lahti, R. (2005) Membrane-bound pyrophosphatase of *Thermotoga maritima* requires sodium for activity, *Biochemistry.* 44, 2088-96.

79. Schulz, S., Iglesias-Cans, M., Krah, A., Yildiz, O., Leone, V., Matthies, D., Cook, G. M., Faraldo-Gomez, J. D. & Meier, T. (2013) A new type of Na^+-driven ATP synthase membrane rotor with a two-carboxylate ion-coupling motif, *PLoS biology.* 11, e1001596.

80. Beatrix, B., Bendrat, K., Rospert, S. & Buckel, W. (1990) The biotin-dependent sodium ion pump glutaconyl-CoA decarboxylase from *Fusobacterium nucleatum* (subsp. nucleatum). Comparison with the glutaconyl-CoA decarboxylases from gram-positive bacteria, *Archives of microbiology.* 154, 362-9.

81. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., Gwinn, M., Dougherty, B., Tomb, J. F., Fleischmann, R. D., Richardson, D., Peterson, J., Kerlavage, A. R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M. D., Gocayne, J., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fuji, C., Cotton, M. D., Horst, K., Roberts, K., Hatch, B., Smith, H. O. &

Venter, J. C. (1997) Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*, *Nature.* 390, 580-6.

82. Hase, C. C., Fedorova, N. D., Galperin, M. Y. & Dibrov, P. A. (2001) Sodium ion cycle in bacterial pathogens: evidence from cross-genome comparisons, *Microbiology and molecular biology reviews : MMBR.* 65, 353-70, table of contents.

83. Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B. C. & Herrmann, R. (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*, *Nucleic acids research.* 24, 4420-49.

84. Wodke, J. A., Puchalka, J., Lluch-Senar, M., Marcos, J., Yus, E., Godinho, M., Gutierrez-Gallego, R., dos Santos, V. A., Serrano, L., Klipp, E. & Maier, T. (2013) Dissecting the energy metabolism in *Mycoplasma pneumoniae* through genome-scale metabolic modeling, *Molecular systems biology.* 9, 653.

85. Costa, K. C. & Leigh, J. A. (2014) Metabolic versatility in methanogens, *Current opinion in biotechnology.* 29, 70-5.

86. Kurkin, S., Meuer, J., Koch, J., Hedderich, R. & Albracht, S. P. (2002) The membrane-bound [NiFe]-hydrogenase (Ech) from *Methanosarcina barkeri*: unusual properties of the iron-sulphur clusters, *European journal of biochemistry / FEBS.* 269, 6101-11.

87. Tilis, A., Akhunbaeva, N. M. & Tursunaliev, D. T. (1977) Effect of barochamber hypoxia on erythropoiesis in response to blood loss in inflammation and overheating, *Zdravookhranenie Kirgizii*, 10-6.

88. Hedderich, R., Hamann, N. & Bennati, M. (2005) Heterodisulfide reductase from methanogenic archaea: a new catalytic role for an iron-sulfur cluster, *Biological chemistry.* 386, 961-70.

89. Welte, C. & Deppenmeier, U. (2014) Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens, *Biochim Biophys Acta.* 1837, 1130-47.

90. Gupta, R. S., Naushad, S. & Baker, S. (2015) Phylogenomic analyses and molecular signatures for the class Halobacteria and its two major clades: a proposal for division of the class Halobacteria into an emended order *Halobacteriales* and two new orders, *Haloferacales* ord. nov. and *Natrialbales* ord. nov., containing the novel

families *Haloferacaceae* fam. nov. and *Natrialbaceae* fam. nov, *International journal of systematic and evolutionary microbiology.* 65, 1050-69.

91. Yoshimatsu, K., Iwasaki, T. & Fujiwara, T. (2002) Sequence and electron paramagnetic resonance analyses of nitrate reductase NarGH from a denitrifying halophilic euryarchaeote *Haloarcula marismortui*, *Febs Lett.* 516, 145-50.

92. Baricz, A., Cristea, A., Muntean, V., Teodosiu, G., Andrei, A. S., Molnar, I., Alexe, M., Rakosy-Tican, E. & Banciu, H. L. (2015) Culturable diversity of aerobic halophilic archaea (Fam. *Halobacteriaceae*) from hypersaline, meromictic Transylvanian lakes, *Extremophiles : life under extreme conditions.* 19, 525-37.

93. Rascovan, N., Maldonado, J., Vazquez, M. P. & Eugenia Farias, M. (2015) Metagenomic study of red biofilms from Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea, *The ISME journal*.

94. Scharf, B., Wittenberg, R. & Engelhard, M. (1997) Electron transfer proteins from the haloalkaliphilic archaeon *Natronobacterium pharaonis*: possible components of the respiratory chain include cytochrome *bc* and a terminal oxidase cytochrome ba_3, *Biochemistry.* 36, 4471-9.

95. Amenabar, M. J., Flores, P. A., Pugin, B., Boehmwald, F. A. & Blamey, J. M. (2013) Archaeal diversity from hydrothermal systems of Deception Island, Antarctica, *Polar Biol.* 36, 373-380.

96. Sapra, R., Verhagen, M. F. & Adams, M. W. (2000) Purification and characterization of a membrane-bound hydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*, *Journal of bacteriology.* 182, 3423-8.

97. Choi, A. R., Kim, M. S., Kang, S. G. & Lee, H. S. (2016) Dimethyl sulfoxide reduction by a hyperhermophilic archaeon *Thermococcus onnurineus* NA1 via a cysteine-cystine redox shuttle, *Journal of microbiology.* 54, 31-8.

98. Thurl, S., Buhrow, I. & Schafer, W. (1985) Quinones from Archaebacteria, I. New types of menaquinones from the thermophilic archaebacterium *Thermoproteus tenax*, *Biological chemistry Hoppe-Seyler.* 366, 1079-83.

99. Thurl, S., Witke, W., Buhrow, I. & Schafer, W. (1986) Quinones from archaebacteria, II. Different types of quinones from sulphur-dependent archaebacteria, *Biological chemistry Hoppe-Seyler.* 367, 191-7.

100. Collins, M. D. & Langworthy, T. A. (1983) Respiratory quinone composition of some acidophilic bacteria, *Systematic and applied microbiology.* 4, 295-304.

101. Ishikawa, R., Ishido, Y., Tachikawa, A., Kawasaki, H., Matsuzawa, H. & Wakagi, T. (2002) *Aeropyrum pernix* K1, a strictly aerobic and hyperthermophilic archaeon, has two terminal oxidases, cytochrome ba_3 and cytochrome aa_3, *Archives of microbiology.* 179, 42-49.

102. Kabashima, Y. & Sakamoto, J. (2011) Purification and biochemical properties of a cytochrome *bc* complex from the aerobic hyperthermophilic archaeon *Aeropyrum pernix*, *BMC microbiology.* 11, 52.

103. Fielding, A. J., Parey, K., Ermler, U., Scheller, S., Jaun, B. & Bennati, M. (2013) Advanced electron paramagnetic resonance on the catalytic iron-sulfur cluster bound to the CCG domain of heterodisulfide reductase and succinate: quinone reductase, *Journal of biological inorganic chemistry : JBIC : a publication of the Society of Biological Inorganic Chemistry.* 18, 905-15.

104. Wakao, H., Wakagi, T. & Oshima, T. (1987) Purification and properties of NADH dehydrogenase from a thermoacidophilic archaebacterium, *Sulfolobus acidocaldarius*, *Journal of biochemistry.* 102, 255-62.

105. Brito, J. A., Sousa, F. L., Stelter, M., Bandeiras, T. M., Vonrhein, C., Teixeira, M., Pereira, M. M. & Archer, M. (2009) Structural and functional insights into sulfide:quinone oxidoreductase, *Biochemistry.* 48, 5613-22.

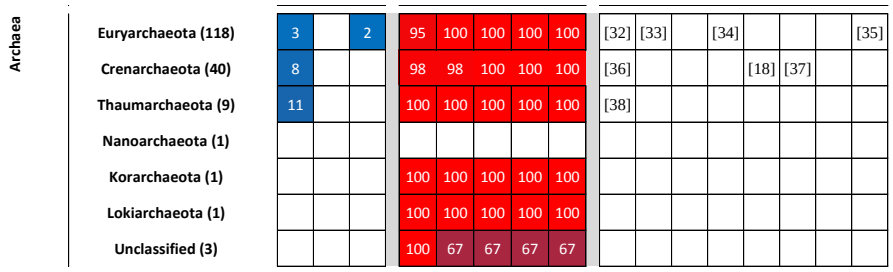
106. Auernik, K. S. & Kelly, R. M. (2008) Identification of components of electron transport chains in the extremely thermoacidophilic crenarchaeon *Metallosphaera sedula* through iron and sulfur compound oxidation transcriptomes, *Applied and environmental microbiology.* 74, 7723-32.

107. Iwasaki, T. (2010) Iron-sulfur world in aerobic and hyperthermoacidophilic archaea Sulfolobus, *Archaea.* 2010.

IV.1.7 Supplementary Material

			Quinones biosynthesis pathway			Isoprenoid tail biosynthesis pathway					According to litterature							
		Phylum (no. of species)	MQ	UQ	PQ/Tocopherol	phytyl-PP (Menaquinone/Tocopherol)	hexaprenyl-PP	octaprenyl-PP (Menaquinone)	nonaprenyl-PP (Plastoquinone)	decaprenyl-PP	MQ	UQ	PQ/Tocopherol	MP	CQ	SQ	RQ	TPQ
Eukarya	Animals	Vertebrates (64)		95	100		98	100	100	100		[1]						
		Lancelets (1)		100	100	100	100	100	100	100								
		Ascidians (1)		100	100		100	100	100	100								
		Echinoderms (1)		100	100	100	100	100	100	100								
		Arthropods (28)		100	96	4	100	100	100	100		[2]						
		Nematodes (5)		60	100		100	100	100	100		[3]					[4]	
		Annelids (1)		100	100		100	100	100	100								
		Mollusks (2)		100	100		100	100	100	100								
		Flatworms (1)			100		100	100	100	100							[5]	
		Cnidarians (2)		100	100		100	100	100	100								
		Placozoans (1)		100	100		100	100	100	100								
		Poriferans (1)		100	100		100	100	100	100								
	Plants	Eudicots (27)		7	100	100	100	100	100	100			[6]					
		Monocots (9)			100	100	100	100	100	100			[6]					
		Basal Magnoliophyta (1)			100	100	100	100	100	100								
		Ferns (1)			100	100	100	100	100	100								
		Mosses (1)			100	100	100	100	100	100								
		Green algae (9)		22	100	100	100	100	100	100								
		Red algae (3)		67	100	100	100	100	100	100								
	Fungi	Ascomycetes (75)		99	79	5	100	100	100	100								
		Basidiomycetes (27)		78	4	4	100	100	100	100								
		Microsporidians (5)						20	20									
		Choanoflagellates (1)			100			100	100									
	Protists	Amoebozoa (6)		33	67	67	100	100	100	100							[7]	[8]
		Alveolates (17)			12		100	94	94	94								
		Stramenopiles (5)		40	80	40	80	80	80	80								
		Haptophyta (1)		100		100	100	100	100	100								
		Cryptomonads (1)			100	100	100	100	100	100								
		Euglenozoa (7)		100			100	100	100	100							[9]	
		Heterolobosea (1)		100	100		100	100	100	100								
		Parabasalids (1)						100	100									
		Diplomonads (1)						100										

Bacteria																	
γ-proteobacteria E. (105)	76	95	5	25	97	99	99	99	[10]	[11]							
γ-proteobacteria O. (220)	28	86	55	17	98	98	98	98									
β-proteobacteria (130)	1	93	64	13	96	96	96	96		[12]							
ε-proteobacteria (40)	100			15	100	100	100	100	[13]								[14]
δ-proteobacteria (58)	60	2	29	41	98	98	98	98	[15]								
α-proteobacteria (196)		70	41	23	99	99	99	99		[16]						[17]	
Firmicutes Bacilli (177)	11	5	7	6	98	99	99	99	[18]								
Firmicutes Clostridia (120)	5			33	98	98	98	98	[19]								
Firmicutes Others (16)				13	94	100	100	100									
Tenericutes (65)				2	3	3	3	3									
Actinobacteria (225)	30	19	36	49	97	100	100	100	[20]	[21]							
Chlamydiae (13)	31				54	100	100	100	[22]								
Verrucomicrobia (6)	17			17	100	100	100	100	[23]								
Spirochaetes (40)	10			5	88	95	95	93									
Acidobacteria (8)	13	25	63	75	100	100	100	100	[24]								
Fibrobacteres (1)				100	100	100	100	100									
Elusimicrobia (3)				33	100	100	100	100									
Fusobacteria (7)				29	71	100	100	100									
Gemmatimonadetes (2)	100		100		100	100	100	100	[25]								
Synergistetes (5)					80	80	80	80									
Planctomycetes (7)	57		14		100	100	100	100	[26]								
Cyanobacteria (46)		74	74	98	100	100	100	100		[27]	[28]						
Bacteroidetes (108)	23		55	18	98	98	98	98	[29]								
Chlorobi (12)	83		8	92	100	100	100	100									
Chloroflexi (14)	7		29	71	93	93	93	93									
Deinococcus-Thermus (17)	41	29	6		100	100	100	100									
Aquificae (10)	80			70	100	100	100	100									
Thermotogae (18)				11	89	100	100	83									
Armatimonadetes (2)			100	50	100	100	100	100									
Caldiserica (1)				100	100	100	100	100									
Chrysiogenetes (1)					100	100	100	100									
Deferribacteres (4)	100				100	100	100	100	[30]								
Dictyoglomi (2)				100	100	100	100	100									
Nitrospirae (6)	17			50	100	100	100	100									
Thermodesulfobacteria (3)	100			33	100	100	100	100	[31]								
Unclassified (9)	11			22	33	44	44	44									



Supplementary Figure IV.1.1: Taxonomic distribution of enzymes involved in the quinones rings and isoprepenoid tails biosynthetic pathways, and quinones types described in literature. MQ – Menaquinone, UQ – Ubiquinone, PQ – Plastoquinone, MP – Methanophenazine, CQ – Caldariellaquinone, SQ – Sulfolobusquinone, RQ – Rhodoquinone, TPQ – Thermoplasmaquinone.

References in Supplementary Figure IV.1.1:

1. Diplock, A. T. & Haslewood, G. A. (1967) The ubiquinone content of animal tissues. A survey of the occurrence of ubiquinone in vertebrates, *Biochem J.* 104, 1004-10.

2. Heller, J., Szarkowska, L. & Michalek, H. (1960) Ubiquinone (coenzyme Q) in insects, *Nature.* 188, 491.

3. Sato, M. & Ozawa, H. (1969) Occurrence of ubiquinone and rhodoquinone in parasitic nematodes, *Metastrangylus elongatus* and *Ascaris lumbricoides* var. suis, *J Biochem.* 65, 861-7.

4. Inaoka, D. K., Shiba, T., Sato, D., Balogun, E. O., Sasaki, T., Nagahama, M., Oda, M., Matsuoka, S., Ohmori, J., Honma, T., Inoue, M., Kita, K. & Harada, S. (2015) Structural Insights into the Molecular Design of Flutolanil Derivatives Targeted for Fumarate Respiration of Parasite Mitochondria, *Int J Mol Sci.* 16, 15287-15308.

5. Shiobara, Y., Harada, C., Shiota, T., Sakamoto, K., Kita, K., Tanaka, S., Tabata, K., Sekie, K., Yamamoto, Y. & Sugiyama, T. (2015) Knockdown of the coenzyme Q synthesis gene Smed-dlp1 affects planarian regeneration and tissue homeostasis, *Redox Biol.* 6, 599-606.

6. Barr, R. & Crane, F. L. (1967) Comparative Studies on Plastoquinones. III. Distribution of Plastoquinones in Higher Plants, *Plant Physiol.* 42, 1255-63.

7. Stairs, C. W., Eme, L., Brown, M. W., Mutsaers, C., Susko, E., Dellaire, G., Soanes, D. M., van der Giezen, M. & Roger, A. J. (2014) A SUF Fe-S Cluster Biogenesis System in the Mitochondrion-Related Organelles of the Anaerobic Protist *Pygsuia*, *Curr Biol.* 24, 1176-1186.

8. Nandi, N., Bera, T., Kumar, S., Purkait, B., Kumar, A. & Das, P. (2011) Involvement of thermoplasmaquinone-7 in transplasma membrane electron transport of *Entamoeba histolytica* trophozoites: a key molecule for future rational chemotherapeutic drug designing, *J Bioenerg Biomembr.* 43, 203-215.

9. Castro-Guerrero, N. A., Jasso-Chavez, R. & Moreno-Sanchez, R. (2005) Physiological role of rhodoquinone in *Euglena gracilis* mitochondria, *BBA-Bioenergetics.* 1710, 113-121.

10. Nishijima, M., Adachi, K., Katsuta, A., Shizuri, Y. & Yamasato, K. (2013) *Endozoicomonas numazuensis* sp nov., a gammaproteobacterium isolated from marine sponges, and emended description of the genus Endozoicomonas Kurahashi and Yokota 2007, *Int J Syst Evol Micr.* 63, 709-714.

11. Fahrbach, M., Kuever, J., Remesch, M., Huber, B. E., Kampfer, P., Dott, W. & Hollender, J. (2008) *Steroidobacter denitrificans* gen. nov., sp nov., a steroidal hormone-degrading gammaproteobacterium, *Int J Syst Evol Micr.* 58, 2215-2223.

12. Kalyuzhnaya, M. G., De Marco, P., Bowerman, S., Pacheco, C. C., Lara, J. C., Lidstrom, M. E. & Chistoserdova, L. (2006) *Methyloversatilis universalis* gen. nov., sp nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates, *Int J Syst Evol Micr.* 56, 2517-2522.

13. Kern, M. & Simon, J. (2009) Electron transport chains and bioenergetics of respiratory nitrogen metabolism in *Wolinella succinogenes* and other Epsilonproteobacteria, *BBA-Bioenergetics.* 1787, 646-656.

14. Collins, M. D. & Fernandez, F. (1984) Menaquinone-6 and Thermoplasmaquinone-6 in *Wolinella-Succinogenes*, *Fems Microbiol Lett.* 22, 273-276.

15. Zacharoff, L., Chan, C. H. & Bond, D. R. (2016) Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of *Geobacter sulfurreducens*, *Bioelectrochemistry.* 107, 7-13.

16. Gonzalez, B., Martinez, S., Chavez, J. L., Lee, S., Castro, N. A., Dominguez, M. A., Gomez, S., Contreras, M. L., Kennedy, C. & Escamilla, J. E. (2006) Respiratory system of *Gluconacetobacter diazotrophicus* PAL5. Evidence for a cyanide-sensitive cytochrome bb and cyanide-resistant cytochrome ba quinol oxidases, *Biochim Biophys Acta.* 1757, 1614-22.

17. Maroti, A., Wraight, C. A. & Maroti, P. (2015) Protonated Rhodosemiquinone at the Q(B) Binding Site of the M265IT Mutant Reaction Center of Photosynthetic Bacterium *Rhodobacter sphaeroides*, *Biochemistry-Us.* 54, 2095-2103.

18. Collins, M. D. & Langworthy, T. A. (1983) Respiratory Quinone Composition of Some Acidophilic Bacteria, *Syst Appl Microbiol.* 4, 295-304.

19. Gottwald, M., Andreesen, J. R., LeGall, J. & Ljungdahl, L. G. (1975) Presence of cytochrome and menaquinone in *Clostridium formicoaceticum* and *Clostridium thermoaceticum*, *J Bacteriol.* 122, 325-8.

20. Falentin, H., Deutsch, S. M., Jan, G., Loux, V., Thierry, A., Parayre, S., Maillard, M. B., Dherbecourt, J., Cousin, F. J., Jardin, J., Siguier, P., Couloux, A., Barbe, V., Vacherie, B., Wincker, P., Gibrat, J. F., Gaillardin, C. & Lortal, S. (2010) The complete genome of *Propionibacterium freudenreichii* CIRM-BIA1, a hardy actinobacterium with food and probiotic applications, *PLoS One.* 5, e11748.

21. Monteiro-Vitorello, C. B., Camargo, L. E., Van Sluys, M. A., Kitajima, J. P., Truffi, D., do Amaral, A. M., Harakava, R., de Oliveira, J. C., Wood, D., de Oliveira, M. C., Miyaki, C., Takita, M. A., da Silva, A. C., Furlan, L. R., Carraro, D. M., Camarotte, G., Almeida, N. F., Jr., Carrer, H., Coutinho, L. L., El-Dorry, H. A., Ferro, M. I., Gagliardi, P. R., Giglioti, E., Goldman, M. H., Goldman, G. H., Kimura, E. T., Ferro, E. S., Kuramae, E. E., Lemos, E. G., Lemos, M. V., Mauro, S. M., Machado, M. A., Marino, C. L., Menck, C. F., Nunes, L. R., Oliveira, R. C., Pereira, G. G., Siqueira, W., de Souza, A. A., Tsai, S. M., Zanca, A. S., Simpson, A. J., Brumbley, S. M. & Setubal, J. C. (2004) The genome sequence of the gram-positive sugarcane pathogen *Leifsonia xyli* subsp. xyli, *Mol Plant Microbe Interact.* 17, 827-36.

22. Barta, M. L., Thomas, K., Yuan, H. L., Lovell, S., Battaile, K. P., Schramm, V. L. & Hefty, P. S. (2014) Structural and Biochemical Characterization of *Chlamydia*

trachomatis Hypothetical Protein CT263 Supports That Menaquinone Synthesis Occurs through the Futalosine Pathway, *J Biol Chem.* 289, 32214-32229.

23. Yoon, J., Matsuo, Y., Matsuda, S., Adachi, K., Kasai, H. & Yokota, A. (2007) Rubritalea spongiae sp nov and *Rubritalea tangerina* sp nov., two carotenoid- and squalene-producing marine bacteria of the family Verrucomicrobiaceae within the phylum 'Verrucomicrobia', isolated from marine animals, *Int J Syst Evol Micr.* 57, 2337-2343.

24. Kulichevskaya, I. S., Kostina, L. A., Valaskova, V., Rijpstra, W. I., Damste, J. S., de Boer, W. & Dedysh, S. N. (2012) *Acidicapsa borealis* gen. nov., sp. nov. and *Acidicapsa ligni* sp. nov., subdivision 1 Acidobacteria from Sphagnum peat and decaying wood, *Int J Syst Evol Microbiol.* 62, 1512-20.

25. Zhang, H., Sekiguchi, Y., Hanada, S., Hugenholtz, P., Kim, H., Kamagata, Y. & Nakamura, K. (2003) *Gemmatimonas aurantiaca* gen. nov., sp. nov., a gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov, *Int J Syst Evol Microbiol.* 53, 1155-63.

26. Clum, A., Tindall, B. J., Sikorski, J., Ivanova, N., Mavrommatis, K., Lucas, S., Glavina, T., Del, R., Nolan, M., Chen, F., Tice, H., Pitluck, S., Cheng, J. F., Chertkov, O., Brettin, T., Han, C., Detter, J. C., Kuske, C., Bruce, D., Goodwin, L., Ovchinikova, G., Pati, A., Mikhailova, N., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. D., Chain, P., Rohde, M., Goker, M., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C., Klenk, H. P. & Lapidus, A. (2009) Complete genome sequence of *Pirellula staleyi* type strain (ATCC 27377), *Stand Genomic Sci.* 1, 308-16.

27. Dufresne, A., Salanoubat, M., Partensky, F., Artiguenave, F., Axmann, I. M., Barbe, V., Duprat, S., Galperin, M. Y., Koonin, E. V., Le Gall, F., Makarova, K. S., Ostrowski, M., Oztas, S., Robert, C., Rogozin, I. B., Scanlan, D. J., Tandeau de Marsac, N., Weissenbach, J., Wincker, P., Wolf, Y. I. & Hess, W. R. (2003) Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome, *Proceedings of the National Academy of Sciences of the United States of America.* 100, 10020-5.

28. Pfaff, C., Glindemann, N., Gruber, J., Frentzen, M. & Sadre, R. (2014) Chorismate Pyruvate-Lyase and 4-Hydroxy-3-solanesylbenzoate Decarboxylase Are Required for Plastoquinone Biosynthesis in the Cyanobacterium *Synechocystis* sp PCC6803, *J Biol Chem.* 289, 2675-2686.

29. Mukherjee, S., Lapidus, A., Shapiro, N., Cheng, J. F., Han, J., Reddy, T., Huntemann, M., Ivanova, N., Mikhailova, N., Chen, A., Palaniappan, K., Spring, S., Goker, M., Markowitz, V., Woyke, T., Tindall, B. J., Klenk, H. P., Kyrpides, N. C. & Pati, A. (2015) High quality draft genome sequence and analysis of *Pontibacter roseus* type strain SRC-1(T) (DSM 17521(T)) isolated from muddy waters of a drainage system in Chandigarh, India, *Stand Genomic Sci.* 10, 8.

30. Takaki, Y., Shimamura, S., Nakagawa, S., Fukuhara, Y., Horikawa, H., Ankai, A., Harada, T., Hosoyama, A., Oguchi, A., Fukui, S., Fujita, N., Takami, H. & Takai, K. (2010) Bacterial lifestyle in a deep-sea hydrothermal vent chimney revealed by the genome sequence of the thermophilic bacterium *Deferribacter desulfuricans* SSM1, *DNA Res.* 17, 123-37.

31. Anderson, I., Saunders, E., Lapidus, A., Nolan, M., Lucas, S., Tice, H., Del Rio, T. G., Cheng, J. F., Han, C., Tapia, R., Goodwin, L. A., Pitluck, S., Liolios, K., Mavromatis, K., Pagani, I., Ivanova, N., Mikhailova, N., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Jeffries, C. D., Chang, Y. J., Brambilla, E. M., Rohde, M., Spring, S., Goker, M., Detter, J. C., Woyke, T., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C. & Klenk, H. P. (2012) Complete genome sequence of the thermophilic sulfate-reducing ocean bacterium *Thermodesulfatator indicus* type strain (CIR29812(T)), *Standards in Genomic Sciences.* 6, 155-164.

32. von Jan, M., Lapidus, A., Del Rio, T. G., Copeland, A., Tice, H., Cheng, J. F., Lucas, S., Chen, F., Nolan, M., Goodwin, L., Han, C., Pitluck, S., Liolios, K., Ivanova, N., Mavromatis, K., Ovchinnikova, G., Chertkov, O., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y. J., Jeffries, C. D., Saunders, E., Brettin, T., Detter, J. C., Chain, P., Eichinger, K., Huber, H., Spring, S., Rohde, M., Goker, M., Wirth, R., Woyke, T., Bristow, J., Eisen, J. A., Markowitz, V., Hugenholtz, P., Kyrpides, N. C. & Klenk, H. P. (2010) Complete genome sequence of *Archaeoglobus profundus* type strain (AV18(T)), *Standards in Genomic Sciences.* 2, 327-346.

33. Ullmann, E., Tan, T. C., Gundinger, T., Herwig, C., Divne, C. & Spadiut, O. (2014) A novel cytosolic NADH:quinone oxidoreductase from *Methanothermobacter marburgensis*, *Biosci Rep.* 34, e00167.

34. Baumer, S., Ide, T., Jacobi, C., Johann, A., Gottschalk, G. & Deppenmeier, U. (2000) The F420H2 dehydrogenase from *Methanosarcina mazei* is a redox-driven proton pump closely related to NADH dehydrogenases, *J Biol Chem.* 275, 17968-17973.

35. Collins, M. D. (1985) Structure of Thermoplasmaquinone from *Thermoplasma-Acidophilum*, *Fems Microbiol Lett.* 28, 21-23.

36. Thurl, S., Witke, W., Buhrow, I. & Schafer, W. (1986) Quinones from archaebacteria, II. Different types of quinones from sulphur-dependent archaebacteria, *Biol Chem Hoppe Seyler.* 367, 191-7.

37. Thurl, S., Buhrow, I. & Schafer, W. (1985) Quinones from Archaebacteria, I. New types of menaquinones from the thermophilic archaebacterium *Thermoproteus tenax*, *Biol Chem Hoppe Seyler.* 366, 1079-83.

38. Elling, F. J., Becker, K. W., Konneke, M., Schroder, J. M., Kellermann, M. Y., Thomm, M. & Hinrichs, K. U. (2015) Respiratory quinones in Archaea: phylogenetic distribution and application as biomarkers in the marine environment, *Environ Microbiol*.

Part III

Final discussion

Chapter V

Final discussion

Chapter V – Final discussion

Chapter V.1 – Final discussion

V.1.1 Acquisition of energy is central to life

Organisms convert most chemical energy into the form of a transmembrane difference in electrochemical potential, used for ATP synthesis, solute transport and motility, which is determined by the global operation of cycles of different ions, namely H^+ and Na^+. A variety of membrane-bound respiratory complexes are involved in both cycles or exclusively in only one.

We brought a novel, broad and integrated perspective of membrane-bound respiratory enzymes, by evaluating the distribution of 26 quinone reductases, 5 quinol:electron carriers oxidoreductases, 18 terminal electron acceptor reductases, quinones or soluble proteins, and redox or decarboxylation driven ion translocases, ATP synthase and transhydrogenase with the rational that the conservation of complexes/enzymes and their combination reflects the several energetic metabolisms of living systems (see Chapter IV). Included in the membrane bound respiratory enzymes are Complex I (see Chapter II) and NDH-2 (see Chapter III). These proteins have NAD(P)H:quinone oxidoreductase activity and are electron entry points of respiratory chains, contributing directly (Complex I) and indirectly (NDH-2) to the establishment of the membrane potential.

V.1.2 Energy transduction by Complex I

We investigated the relationships among all members of Complex I family, including Complex I and Complex I-like enzymes, group 4 [NiFe] hydrogenases and energy-converting hydrogenase related complexes (Ehr). We identified the common denominator among these enzymes composed of four subunits homologous to the peripheral subunits NuoB and D and to the membrane subunits NuoH and a Na^+/H^+ antiporter-like protein (see section II.1, [1]). This antiporter is characterized by the presence of a long amphipathic α-helix (HL) at the C-terminus (see section II.2, [2]).

We proposed that Na^+/H^+ antiporter activity reflects a possible evolutionary origin of Complex I [1]. It is well established that complex I has subunits homologous to subunits of Mrp Na^+/H^+ antiporters, soluble NAD^+ reducing hydrogenases and group 4 membrane-bound [NiFe] hydrogenases [3-5]. We hypothesized that the precursor of complex I, membrane bound hydrogenases and Ehrs was the result of the association of a soluble enzyme performing a thermodynamic favorable redox reaction with a transmembrane protein performing passive or secondary charge transport [1]. In this way, we put forward the idea that a soluble redox enzyme performing electron donor: H^+ oxidoreductase activity, which gave rise to the precursor of the peripheral arm of group 4 membrane-bound [NiFe] hydrogenases, would have associated with a Na^+/H^+ antiporter to allow energy coupling. This antiporter in the new established complex would work in the reverse direction of that performed when isolated and in this way would contribute to the establishment of the transmembrane

difference of electrochemical potential. Specifically, we suggested that this complex was the precursor of group 4 membrane-bound [NiFe] hydrogenases and later of Ehr and Complex I by loss of the [NiFe] catalytic center [1] (Figure V.1.1).

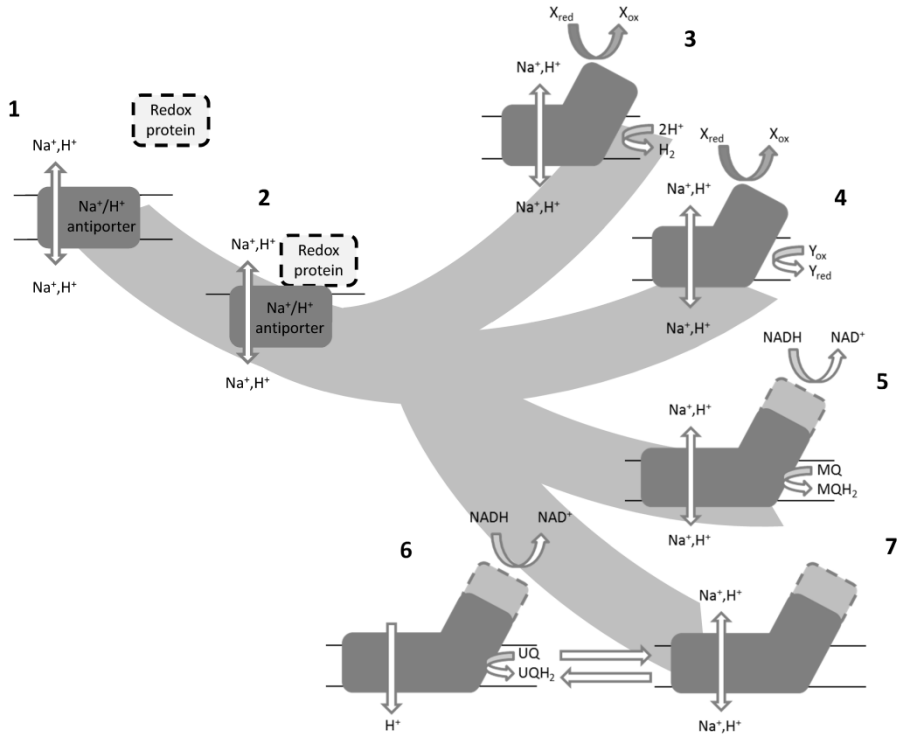


Figure V.1.1. Proposal of an evolutionary scenario for complex I. A soluble redox enzyme performing electron donor:H^+ oxidoreductase activity would have associated with a Na^+/H^+ antiporter in order to allow energy coupling and generation of transmembrane difference of electrochemical potential (1 and 2). This complex would give rise to group 4 membrane-bound [NiFe] hydrogenases (3), energy-converting hydrogenase related complexes (4) and complex I (5 and 6). Menaquinone-reducing complexes I (5) need to perform Na^+/H^+ antiporter activity in order to achieve the $4H^+/2e^-$ stoichiometry. Ubiquinone-reducing complexes I transport solely protons when coupled to the catalytic reaction (6) but could have kept Na^+/H^+ antiporter activity independently of catalytic activity, in its deactive form, for example (7). Figure adapted from [6].

Complex I is one of the most complex respiratory enzymes present in the mitochondria of Eukarya and in the respiratory chains of Bacteria. Several reports suggested that Complex I is involved in proton and sodium bioenergetics, either by using proton or sodium as coupling ion, or by performing Na^+/H^+ antiporter activity (see section I.2).

Na^+/H^+ antiporter activity was shown to occur by the action of Complexes I in Bacteria and Eukarya. However, the physiological relevance of Na^+ transport in these organisms may be different; in the first case it may be a requirement in order to couple the redox reaction to the translocation of 4 ions and in the second case, the antiporter activity may be under tight regulation and only operative in certain conditions, as a reflex of the sophistication of eukaryotic Complex I. So far, two bacterial complexes I were shown to perform Na^+/H^+ antiporter activity during catalytic turnover [7-9] (see sections II.3 and II.4). We hypothesized that the need for Na^+ transport by Complex I is related to the available redox energy, ΔE, defined by the type of quinone used by the organism, menaquinone or ubiquinone. The two bacteria whose complexes I were observed to have Na^+/H^+ antiporter activity use a low potential quinone, menaquinone, as electron acceptor. The menaquinone-reducing Complexes I would need an extra demand of energy to reach the accepted stoichiometry of $4H^+/2e^-$. This extra energy could be achieved by dissipation of the Na^+ electrochemical potential. In ubiquinone dependent complexes I, enough ΔE is already available by the redox reaction itself to transport 4 protons across the membrane.

Therefore, we suggested both Bacterial and Eukarya Complexes I that use ubiquinone would be specialized to transport solely protons when coupled to the catalytic reaction. Moreover, in ubiquinone dependent Complexes I, Na^+ transport activity could have been kept independently of catalytic activity and could be activated under specific circumstances. As suggested by Hirst and coworkers [10], who observed Na^+/H^+ antiporter activity by the deactive form of Complex I from *B. taurus* mitochondria, such a specific circumstance could be an ischemia-reperfusion event. The authors discussed that complex I driven transmembrane fluxes of Na^+ and H^+ may influence intra-mitochondrial Ca^{2+} homeostasis, which is related to important cell functions, including apoptosis [10, 11]. The results obtained for the bovine enzyme suggest Na^+/H^+ antiporter activity was maintained in ubiquinone-reducing Complexes I, but under different circumstances. This may reflect an evolutionary step forward in the regulation and operative modes of the complex (Figure V.1.1) [6].

As far as we know, H^+ and Na^+ ions have been described to be translocated by members of all types of ion pumping enzymes, namely Na^+-NQR [12], HCO [13], ATPases [14], Na^+-mPPases [15], aspartate transporter GltPh [16] and Na^+/H^+ antiporters [17-22]. However, the structural determinants of Na^+ and H^+ binding sites are not clearly and unambiguously distinguishable yet. The Na^+/H^+ antiporter activity by Complex I makes this complex no exception to the rule.

Subunits NuoL, M and N of complex I present structural features similar to the ones observed in *bona fide* antiporters or

transporters [23, 24]. In fact, those subunits contain conserved charged or polar amino acid residues in the middle of the membrane, near to discontinuous or partially unwound transmembrane helices. Conserved glutamate, lysine and histidine residues were suggested to play a key role in ion translocation by Complex I [24]. In order to identify the structural motifs involved on Na^+/H^+ antiporter activity and the type of ion that is translocated in each site of the NuoL, M and N subunits of complex I, we are performing amino acid sequence and phylogenetic analysis of all Na^+/H^+ antiporters-like subunits from Complex I family.

Complex I seems to have a very sophisticated structure for its function, specifically when compared for example to the most common last enzyme of aerobic respiratory chains, HCO, which couples O_2 reduction to proton pumping. The free energy available from the oxidation of, for example, cytochrome *c* for O_2 reduction to water drives the pumping of up to 4 protons across the mitochondrial inner membrane (H^+/e^- stoichiometry of 1, half of the stoichiometry of Complex I) [25]. O_2 reduction and proton translocation take place in the same subunit, which has 50 kDa. This is in sharp contrast to Complex I that has 550 kDa in bacteria or 1 MDa in eukaryotes and in which the catalytic reaction module involves seven subunits, while ion translocation seems to take place in several subunits. In fact, as mentioned before, 4 proton translocation sites were hypothesized in Complex I, 3 of them present in the antiporter like subunits with approximately 50 kDa each. In the two systems, complex I and HCOs,

proton translocation is coupled to redox reactions and thus we find it most intriguing the large sophistication of complex I when comparing to HCOs. Why would cells invest so much energy building up such a complex machine when simpler solutions exist? We believe that complex I has a more complex operative mechanism than a simply "coupling" of a redox reaction to ion transport [6]. For instance, the long peripheral arm may work as a *mechanical lever* for energy coupling, Complex I may be assembled with different antiporter-like subunits or supernumerary subunits to transport different ions or perform different functions, as the case of cyanobacterial Complex I-like [26, 27].

V.1.3 Phylogenetic diversity and catalytic mechanism of NDH-2

In order to further understand the role of NDH-2 in cellular metabolism and to obtain a deeper insight into the function of these enzymes, the phylogenetic distribution and conserved structural features of this enzyme were investigated using bioinformatics analysis. NDH-2s are part of the two-Dinucleotide Binding Domains Flavoproteins (tDBDF) superfamily and involved in respiratory chains of organisms belonging to the three domains of life. These are membrane associated enzymes, which do not perform charge translocation and thus contribute indirectly (reduce quinone) to the establishment and maintenance of the transmembrane difference of electrochemical potential.

We obtained the first database of NDH-2 family and identified the sulfide:quinone oxidoreductases (SQR) as the closest related enzymes of the tDBDF. We noticed the absence of genes coding for NDH-2 in anaerobic phyla and we observed three different origins for the eukaryotic sequences and possible lateral gene transfer among prokaryotes in NDH-2 family distribution. Besides eukaryotic NDH-2 from mitochondrial and chloroplastidial origins we observed a third origin that includes enzymes from mammals, namely that from *H. sapiens*, previously assigned as apoptosis-inducing factor-homologous mitochondrion-associated inducer of death (AIF-M2) and suggested to be another apoptosis inducing factor (AIF) (see section III.1). Moreover, NADH:quinone oxidoreductase activity was observed to be performed by AIF-M2 from *H. sapiens* which corroborate our results that AIF-M2 belongs to the NDH-2 family [28]. The nonexistence of NDH-2s in mammals and the fact that these are the only NADH:quinone oxidoreductase expressed in some pathogenic microorganisms were reasons invoked to propose NDH-2 as an attractive target for rational design of specific drugs. Our observations question this idea.

In addition we propose that NDH-2 could appear as the solution to regenerate NAD^+ without contributing directly to the membrane potential which could be advantageous under certain conditions (see section III.1).

We identified conserved amino acid residues, sequence motifs and structural elements which allowed us to systematize several properties of the NDH-2 family, namely quinone binding site motifs, putative regulatory sites, specificity for NADH *vs* NADPH and a series of conserved amino acids that can form a proton pathway in each dinucleotide binding domain. The latter may conduct protons from the bulk to the quinone binding pocket necessary for its protonation. In addition, the different amino acid arrangement of the proton pathway from the 1st dinucleotide binding domain may define the type of quinone used by NDH-2s (see section III.2). These conserved common elements of NDH-2 suggest that NDH-2 have independent substrates binding sites in agreement with the structure obtained by Feng and co-workers that showed distinct binding sites for NADH and quinone at the *re*- and *si*-side of the flavin, respectively [29].

We also proposed a 'proton gate' key role for the conserved glutamate localized near the FAD cofactor and NAD(P)H binding site. We hypothesized that this glutamate residue switches proton bonds between residues belonging to the proton pathway at the 2nd dinucleotide domain and NAD(P)H, which allows protons to be transferred to the quinone binding pocket (see section III.2). In summary we proposed a catalytic mechanism of NDH-2 by investigating the common elements of all NDH-2, based on the rational that the conservation of such elements reflects their structural/functional importance. This is the first proposal of a catalytic mechanism for NDH-2s that may also give insights in the catalytic mechanism for tDBDF members. In order to confirm these

hypotheses and to obtain a deeper insight into the function of NDH-2, biochemical and biophysical analysis using point mutations in conserved amino acids of the NDH-2 are being performed at the host laboratory.

V.1.4 Final remarks

Our results addressed the diversity of respiratory membrane enzymes, especially the type I and type II NADH:quinone oxidoreductases. We searched for common features that characterize theses enzymes and we found the minimal functional elements relevant for their operation. These results contributed to a broader knowledge of Complex I and NDH-2 families and provide new perspectives and discussions for the Bioenergetics community.

V.1.5 References

1. Marreiros, B. C., Batista, A. P., Duarte, A. M. & Pereira, M. M. (2013) A missing link between Complex I and group 4 membrane-bound [NiFe] hydrogenases, *Biochimica et biophysica acta.* 1827, 198-209.

2. Batista, A. P., Marreiros, B. C. & Pereira, M. M. (2013) The antiporter-like subunit constituent of the universal adaptor of complex I, group 4 membrane-bound [NiFe]-hydrogenases and related complexes, *Biological chemistry.* 394, 659-66.

3. Pilkington, S. J., Skehel, J. M., Gennis, R. B. & Walker, J. E. (1991) Relationship between mitochondrial NADH-ubiquinone reductase and a bacterial NAD-reducing hydrogenase, *Biochemistry.* 30, 2166-75.

4. Friedrich, T. & Weiss, H. (1997) Modular evolution of the respiratory NADH:ubiquinone oxidoreductase and the origin of its modules, *J Theor Biol.* 187, 529-40.

5. Mathiesen, C. & Hagerhall, C. (2003) The 'antiporter module' of respiratory chain complex I includes the MrpC/NuoK subunit -- a revision of the modular evolution scheme, *FEBS Lett.* 549, 7-13.

6. Castro, P. J., Silva, A. F., Marreiros, B. C., Batista, A. P. & Pereira, M. M. (2015) Respiratory complex I: A dual relation with H and Na?, *Biochimica et biophysica acta*.

7. Batista, A. P., Marreiros, B. C. & Pereira, M. M. (2011) Decoupling of the catalytic and transport activities of complex I from *Rhodothermus marinus* by sodium/proton antiporter inhibitor, *ACS chemical biology.* 6, 477-83.

8. Batista, A. P., Fernandes, A. S., Louro, R. O., Steuber, J. & Pereira, M. M. (2010) Energy conservation by *Rhodothermus marinus* respiratory complex I, *Biochimica et biophysica acta.* 1797, 509-15.

9. Batista, A. P. & Pereira, M. M. (2011) Sodium influence on energy transduction by complexes I from *Escherichia coli* and *Paracoccus denitrificans*, *Biochimica et biophysica acta.* 1807, 286-92.

10. Roberts, P. G. & Hirst, J. (2012) The deactive form of respiratory complex I from mammalian mitochondria is a Na+/H+ antiporter, *J Biol Chem.* 287, 34743-51.

11. Contreras, L., Drago, I., Zampese, E. & Pozzan, T. (2010) Mitochondria: the calcium connection, *Biochimica et biophysica acta.* 1797, 607-18.

12. Juarez, O. & Barquera, B. (2012) Insights into the mechanism of electron transfer and sodium translocation of the Na(+)-pumping NADH:quinone oxidoreductase, *Biochimica et biophysica acta.* 1817, 1823-32.

13. Muntyan, M. S., Cherepanov, D. A., Malinen, A. M., Bloch, D. A., Sorokin, D. Y., Severina, II, Ivashina, T. V., Lahti, R., Muyzer, G. & Skulachev, V. P. (2015) Cytochrome cbb3 of *Thioalkalivibrio* is a Na+-pumping cytochrome oxidase, *Proceedings of the National Academy of Sciences of the United States of America.* 112, 7695-700.

14. Mayer, F., Leone, V., Langer, J. D., Faraldo-Gomez, J. D. & Muller, V. (2012) A c subunit with four transmembrane helices and one ion (Na+)-binding site in an

archaeal ATP synthase: implications for c ring function and structure, *J Biol Chem.* 287, 39327-37.

15. Luoto, H. H., Belogurov, G. A., Baykov, A. A., Lahti, R. & Malinen, A. M. (2011) Na+-translocating membrane pyrophosphatases are widespread in the microbial world and evolutionarily precede H+-translocating pyrophosphatases, *J Biol Chem.* 286, 21633-42.

16. Boudker, O., Ryan, R. M., Yernool, D., Shimamoto, K. & Gouaux, E. (2007) Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter, *Nature.* 445, 387-93.

17. Inoue, H., Noumi, T., Tsuchiya, T. & Kanazawa, H. (1995) Essential aspartic acid residues, Asp-133, Asp-163 and Asp-164, in the transmembrane helices of a Na+/H+ antiporter (NhaA) from *Escherichia coli*, *FEBS Lett.* 363, 264-8.

18. Padan, E., Venturi, M., Gerchman, Y. & Dover, N. (2001) Na(+)/H(+) antiporters, *Biochimica et biophysica acta.* 1505, 144-57.

19. Jiang, J., Wang, L., Zou, Y., Lu, W., Zhao, B., Zhang, B., Yang, S. & Yang, L. (2013) Identification of important charged residues for alkali cation exchange or pH regulation of NhaH, a Na(+)/H(+) antiporter of *Halobacillus dabanensis*, *Biochimica et biophysica acta.* 1828, 997-1003.

20. Hellmer, J., Teubner, A. & Zeilinger, C. (2003) Conserved arginine and aspartate residues are critical for function of MjNhaP1, a Na+/H+ antiporter of M. jannaschii, *FEBS Lett.* 542, 32-6.

21. Ostroumov, E., Dzioba, J., Loewen, P. C. & Dibrov, P. (2002) Asp(344) and Thr(345) are critical for cation exchange mediated by NhaD, Na(+)/H(+) antiporter of *Vibrio cholerae*, *Biochimica et biophysica acta.* 1564, 99-106.

22. Dibrov, P., Young, P. G. & Fliegel, L. (1998) Functional analysis of amino acid residues essential for activity in the Na+/H+ exchanger of fission yeast, *Biochemistry.* 37, 8282-8.

23. Screpanti, E. & Hunte, C. (2007) Discontinuous membrane helices in transport proteins and their correlation with function, *J Struct Biol.* 159, 261-7.

24. Efremov, R. G. & Sazanov, L. A. (2011) Structure of the membrane domain of respiratory complex I, *Nature.* 476, 414-20.

25. Wikstrom, M. (1989) Identification of the electron transfers in cytochrome oxidase that are coupled to proton-pumping, *Nature.* 338, 776-8.

26. Hu, P., Lv, J., Fu, P. & Hualing, M. (2013) Enzymatic characterization of an active NDH complex from *Thermosynechococcus elongatus*, *FEBS Lett.* 587, 2340-5.

27. Peltier, G., Aro, E. M. & Shikanai, T. (2015) NDH-1 and NDH-2 Plastoquinone Reductases in Oxygenic Photosynthesis, *Annual review of plant biology*.

28. Elguindy, M. M. & Nakamaru-Ogiso, E. (2015) Apoptosis Inducing Factor (AIF) and its Family Member, AMID, are Rotenone-Sensitive NADH:ubiquinone Oxidoreductases (NDH-2), *J Biol Chem.* 290, 20815-20826.

29. Feng, Y., Li, W., Li, J., Wang, J., Ge, J., Xu, D., Liu, Y., Wu, K., Zeng, Q., Wu, J. W., Tian, C., Zhou, B. & Yang, M. (2012) Structural insight into the type-II mitochondrial NADH dehydrogenases, *Nature.* 491, 478-82.